

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

CC The present sequence represents a monocyte mature differentiation
CC factor (MMD) which maintains the life of macrophages for long periods
CC in liquid culture. MMD can be used as an anti-cancer agent, an immune
CC activator and to treat infectious diseases.

XX Sequence 381 AA;

Query Match 100.0%; Score 2115; DB 18; Length 381;
Best Local Similarity 100.0%; Pred. No. 5.1e-154;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRALSTCPACHCPLKAPKAPGVGLVADGGCCCKVCAKOL 60
DB 1 MSSRIARALALVYTLHLTRALSTCPACHCPLKAPKAPGVGLVADGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKCNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLKCNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVTGOCCEWVCEDESIKDPMEDQDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVTGOCCEWVCEDESIKDPMEDQDGLLG 180
QY 181 KELGFDASEVELTRNNELIANGKSSSLKRLPVGMERILLYNPLQOKCIYQTTSMQCS 240
DB 181 KELGFDASEVELTRNNELIANGKSSSLKRLPVGMERILLYNPLQOKCIYQTTSMQCS 240
QY 241 KTCGTGISTRTVNDNPECRIVKTRICEVAPCGQPVYSSLKKGKSKTKKSPPEVAFY 300
DB 241 KTCGTGISTRTVNDNPECRIVKTRICEVAPCGQPVYSSLKKGKSKTKKSPPEVAFY 300
QY 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCDETFESKNVMIQSCKNYNC 360
DB 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCDETFESKNVMIQSCKNYNC 360
QY 361 HANEAAFPYRLFNIDHKFRD 381
DB 361 HANEAAFPYRLFNIDHKFRD 381

RESULT 2
AAB90773
ID AAB90773 standard; Protein; 381 AA.

XX AC AAB90773;

XX DT 15-JUN-2001 (first entry)

XX DE Human shear stress-response protein SEQ ID NO: 46.

XX KW Human; shear stress-response protein; vascular disease;
arteriosclerosis.

XX OS Homo sapiens.

XX PN W0200125427-A1.

XX PD 12-APR-2001.

XX PF 02-OCT-2000; 2000MO-JF06840.

XX PR 01-OCT-1999; 99JP-0280976.

XX PA (RYOW) KYOMA HAKKO KOGYO KK.

XX PA (NOJI/) NOJIMA H.

XX PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;
Kuga T, Sekine S, Nakamura Y, Sugano S;

XX DR WPI; 2001-266308/27.

XX DR N-PSDB; AA02896.

XX PT DNA sequences, proteins encoded by them and antibodies against them

PT useful in diagnosis and treatment of vascular disease caused by
PT arteriosclerosis -
XX
XX Claim 60; Page 345-346; 678pp; Japanese.

CC The present invention provides the protein and coding sequences of a
CC number of human shear stress response proteins. These are useful in the
CC diagnosis, treatment and screening of vascular diseases caused by
CC arteriosclerosis, including heart failure, post-PTCA restenosis and
CC hypertension.

XX SQ Sequence 381 AA;

Query Match 100.0%; Score 2115; DB 22; Length 381;
Best Local Similarity 100.0%; Pred. No. 5.1e-154;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRALSTCPACHCPLKAPKAPGVGLVADGGCCCKVCAKOL 60
DB 1 MSSRIARALALVYTLHLTRALSTCPACHCPLKAPKAPGVGLVADGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKCNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFQPNCKHQ 120
DB 61 NEDCSKTQPCDHTKGLKCNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVTGOCCEWVCEDESIKDPMEDQDGLLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVTGOCCEWVCEDESIKDPMEDQDGLLG 180
QY 181 KELGFDASEVELTRNNELIANGKSSSLKRLPVGMERILLYNPLQOKCIYQTTSMQCS 240
DB 181 KELGFDASEVELTRNNELIANGKSSSLKRLPVGMERILLYNPLQOKCIYQTTSMQCS 240
QY 241 KTCGTGISTRTVNDNPECRIVKTRICEVAPCGQPVYSSLKKGKSKTKKSPPEVAFY 300
DB 241 KTCGTGISTRTVNDNPECRIVKTRICEVAPCGQPVYSSLKKGKSKTKKSPPEVAFY 300
QY 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCDETFESKNVMIQSCKNYNC 360
DB 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCDETFESKNVMIQSCKNYNC 360
QY 361 HANEAAFPYRLFNIDHKFRD 381
DB 361 HANEAAFPYRLFNIDHKFRD 381

RESULT 3
AAU79761
ID AAU79761 standard; Protein; 381 AA.

XX AC AAU79761;

XX DT 30-JUL-2002 (first entry)

XX DE Human Cyr61 protein.

XX KW Human; uterine leiomyoma proliferation; uterine leiomyoma formation;
Cyr61; cytostatic.

XX OS Homo sapiens.

XX PN W0200226193-A2.

XX PD 04-APR-2002.

XX PF 28-SEP-2001; 2001MO-US30783.

XX PR 29-SEP-2000; 2000US-236887P.

XX PA (AMHP) AMERICAN HOME PROD CORP.

XX PI Zhang Z, Sampath D, Zhu Y, Winneker R;

DR	WP1: 2002-383245/41.
DR	N-PSDB; ABK46899.
XX	
PT	Preventing uterine leiomyoma formation or inhibiting proliferation of
PT	uterine leiomyoma in subject, comprises modulating or increasing the
PT	level of Cyr61 in leiomyoma tissue
XX	
PS	Disclosure; Fig 6; 92pp; English.
XX	
CC	The present invention relates to a method of inhibiting proliferation
CC	of uterine leiomyoma or preventing uterine leiomyoma formation. The
CC	method comprises increasing the level of Cyr61 in leiomyoma tissue.
CC	The invention also describes compounds and compositions that stimulate
CC	induction of the Cyr61 gene and compounds that increase Cyr61 activity.
CC	The compositions and the method of the invention are useful for
CC	preventing uterine leiomyoma formation or inhibiting proliferation of
CC	uterine leiomyoma in a subject. The method is particularly useful for
CC	treating or preventing uterine leiomyoma formation, or inhibiting
CC	proliferation of uterine leiomyoma in a subject. The present sequence
CC	represents human Cyr61.
CC	Note: The present sequence shown in Fig 6 is not shown in the correct
CC	sequence order in the figure. The start of the sequence is shown on
CC	page 8/10 of the figures and the rest of the sequence is shown on
CC	page 6/10 of the figures.
CC	
XX	
XX	Sequence 381 AA;
Query Match	100.0%; Score 2115; DB 23; Length 381;
Best Local Similarity	100.0%; Pred. No. 5.1e-154;
Matches 381;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MSSRIARALWVLLHLTRLALSTCPAACCHPLEAPKAGVGLVRDGGCCCKYCAQOL 60
DB	1 MSSRIARALWVLLHLTRLALSTCPAACCHPLEAPKAGVGLVRDGGCCCKYCAQOL 60
QY	61 NEDCSKTOPCDHTKGLKLECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
DB	61 NEDCSKTOPCDHTKGLKLECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
QY	121 CTCIDGAVGCIPLCPQELSLPNIACCPNRLVKVNGCCCEBWWCDEDSIKDMEODGILG 180
DB	121 CTCIDGAVGCIPLCPQELSLPNIACCPNRLVKVNGCCCEBWWCDEDSIKDMEODGILG 180
QY	181 KELGFDASEVELTRNNELIANGKSSSLKRLPVFGMEPRILINPLDQKCIYOTTSWSSCS 240
DB	181 KELGFDASEVELTRNNELIANGKSSSLKRLPVFGMEPRILINPLDQKCIYOTTSWSSCS 240
QY	241 KTCGTGISTRYTNDNPECRIVKETRICEVPRPGQPVYSSLLKGGKCSKTKKSPPEVRYTY 300
DB	241 KTCGTGISTRYTNDNPECRIVKETRICEVPRPGQPVYSSLLKGGKCSKTKKSPPEVRYTY 300
QY	301 AGCLSYVKRYRKYGSCVDGRCMPOLTRTYKMFRCDEGTFSGNNVMIMIOSCKNKNCP 360
DB	301 AGCLSYVKRYRKYGSCVDGRCMPOLTRTYKMFRCDEGTFSGNNVMIMIOSCKNKNCP 360
QY	361 HANEAAFPFYRLFNDIHKFRD. 381
DB	361 HANEAAFPFYRLFNDIHKFRD 381
RESULT 4	
AAE18107	
ID	AAE18107 standard; Protein; 381 AA.
XX	
AC	AAE18107;
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Human connective tissue growth factor-2 (CTGF-2).
KW	Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
KW	ischemia; restenosis; tissue repair; wound healing; congenital defect;
KW	cardiovascular disease; atherosclerosis; heart failure; angina; trauma;

KW	burns; osteoporosis; periodontal disease; liver failure; tranquilizer;
RW	vulnerable; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
KM	gene therapy.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	/note- "Immunogenic epitope"
FT	47..51
FT	/note- "Immunogenic epitope"
FT	59..75
FT	/note- "Immunogenic epitope"
FT	91..119
FT	/note- "Immunogenic epitope"
FT	145..150
FT	/note- "Immunogenic epitope"
FT	164..176
FT	/note- "Immunogenic epitope"
FT	202..208
FT	/note- "Immunogenic epitope"
FT	223..228
FT	/note- "Immunogenic epitope"
FT	239..244
FT	/note- "Immunogenic epitope"
FT	250..257
FT	/note- "Immunogenic epitope"
FT	279..296
FT	/note- "Immunogenic epitope"
FT	307..314
FT	/note- "Immunogenic epitope"
FT	318..323
FT	/note- "Immunogenic epitope"
FT	337..343
FT	/note- "Immunogenic epitope"
XX	
PN	WO200204480-A2.
PD	
XD	17-JAN-2002.
XX	
XP	11-JUL-2001; 2001MO-US21799.
XX	
PR	11-JUL-2000; 2000US-217402P.
PR	18-MAY-2001; 2001US-291642P.
XX	
PA	(HUMA-) HUMAN GENEOME SCI INC.
XX	(TRGE) TRANSGENE SA.
PI	
DR	Li H, Adams MD, Calenda V, Pataciotoli V;
DR	WPI: 2002-171698/22.
XX	N-PSDB: AAD29035.
PT	
PT	stimulating angiogenesis in a mammal preferably human having ischemia
PT	or restenosis or is treated for limb revascularization, by
PT	administering connective tissue growth factor-2 polypeptide or
PT	polynucleotide
XX	
PS	Example 1; Fig 1; 131pp; English.
CC	
CC	The present invention relates to a method for stimulating angiogenesis in
CC	a mammal. The method comprises administering a polynucleotide encoding
CC	connective tissue growth factor-2 (CTGF-2) or an active fragment or its
CC	derivative. The method is useful for stimulating angiogenesis in a mammal
CC	preferably human having ischemia or restenosis or is treated for limb
CC	revascularisation which is leg or arm. The invention is useful for
CC	inhibiting tumor growth, where angiogenesis is utilised for enhancing
CC	the repair of connective and support tissue, promoting the attachment,
CC	fixation and stabilisation of tissue implants and enhancing wound
CC	healing, hence is useful for treating cardiovascular disease e.g.
CC	atherosclerosis, reperfusion injury such as heart failure, angina,
CC	ischemia; and is also used to differentiate, proliferate and attract
CC	cells leading to regeneration of tissues which is utilised to repair
CC	tissues.

CC replace or protect tissue damaged by congenital defects, trauma (burns,
 CC ulcer, etc), age, disease (e.g. osteoporosis, periodontal disease,
 CC liver failure), surgery including cosmetic plastic surgery. The present
 CC sequence is human CTGF-2. CTGF-2 gene is useful in gene therapy.

XX Sequence 381 AA;

Query Match 100.0%; Score 2115; DB 23; Length 381;
 Best Local Similarity 100.0%; Pred. No. 5.1e-154;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVGLVDDGCGCCAKOL 60
 DB 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVGLVDDGCGCCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLKCNFGASSTALKGICRAQSGRCEYNSRITONGESQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLKCNFGASSTALKGICRAQSGRCEYNSRITONGESQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEWVDEDSINDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEWVDEDSINDPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIAGVGSSSLKRLPVFGMEPRILYNPLQOKCIYVOTTSWQCS 240
 DB 181 KELGFDASEVELTRNNELIAGVGSSSLKRLPVFGMEPRILYNPLQOKCIYVOTTSWQCS 240
 QY 241 KTCGTGISTRTVNDNNECRLYKTRICEVRPCQPYSSLSLKKKSKTKKSPKPYRFTY 300
 DB 241 KTCGTGISTRTVNDNNECRLYKTRICEVRPCQPYSSLSLKKKSKTKKSPKPYRFTY 300
 QY 301 AGCLSVKXKTRPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
 DB 301 AGCLSVKXKTRPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
 QY 361 HANEAAFPYRLFNIDHKFRD 381
 DB 361 HANEAAFPYRLFNIDHKFRD 381

RESULT 5
 ID ABB05438 standard; Protein; 381 AA.

XX ABB05438;
 AC ABB05438;
 DT 15-APR-2002 (first entry)

XX Human Cyr61 protein SEQ ID NO:2.

XX Human; Cyr61; breast cancer; sex steroid receptor; cytosolic; promoter;
 KW sex steroid response element; cysteine rich heparin-binding protein;
 KW cell proliferation; heparin binding epidermal growth factor;
 KW epidermal growth factor; basic fibroblastic growth factor.

XX Homo sapiens.

XX W0200198359-A2.

XX 27-DEC-2001.

XX 21-JUN-2001; 2001MO-US19823.

XX 21-JUN-2000; 2000US-213182P.

XX 16-MAY-2001; 2001US-291510P.

XX (AMHP) AMERICAN HOME PROD CORP.

XX Sampath D, Zhang Z, Wineker R;

XX WPI; 2002-147796/19.

XX N-PSDB; ABA93127, ABA93130.

PT Regulation of Cyr61 expression and activity for preventing and
 CC inhibiting breast cancer comprises use of a Cyr61 neutralizing
 PT antibody, an anti-sense oligonucleotide and an antibody which
 XX
 XX
 XX Claim 6; Fig 1; 86pp; English.

CC The present invention describes a method for the prevention or inhibition
 CC of breast cancer cell proliferation. The method comprises administration
 CC of a compound that inhibits the interaction of a sex steroid receptor
 CC with a sex steroid response element of the Cyr61 (cysteine rich heparin-
 CC binding protein) promoter. Cyr61 has cytosolic activity. An antibody (1)
 CC which neutralises Cyr61 can be used to prevent or inhibit breast cancer
 CC cell proliferation by blocking sex steroid induced and growth factor
 CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal,
 CC heparin binding epidermal or basic fibroblastic growth factor. (1) can be
 CC used to diagnose or stage breast cancer where the level of Cyr61 in a
 CC positive/suspect breast cancer cell is compared to the level in a
 CC normal cell, an increase in the level of Cyr61 compared to the level in
 CC normal tissue indicates the presence of breast cancer. The level of
 CC Cyr61 being determined by exposing the tissues to (1), and an increase
 CC in the level of bound antibody by the suspect/positive cell as compared
 CC to the normal tissue indicates the presence of breast cancer. The present
 CC sequence represents the human Cyr61 protein, which is used in the
 CC exemplification of the present invention.

XX Sequence 381 AA;

Query Match 100.0%; Score 2115; DB 23; Length 381;
 Best Local Similarity 100.0%; Pred. No. 5.1e-154;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVGLVDDGCGCCAKOL 60
 DB 1 MSSRIARALALVYTLHLTRLALSTCPAACHCPLKAPGVGLVDDGCGCCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLKCNFGASSTALKGICRAQSGRCEYNSRITONGESQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLKCNFGASSTALKGICRAQSGRCEYNSRITONGESQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEWVDEDSINDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVKTGOCCEWVDEDSINDPMEDDGLLG 180
 QY 181 KELGFDASEVELTRNNELIAGVGSSSLKRLPVFGMEPRILYNPLQOKCIYVOTTSWQCS 240
 DB 181 KELGFDASEVELTRNNELIAGVGSSSLKRLPVFGMEPRILYNPLQOKCIYVOTTSWQCS 240
 QY 241 KTCGTGISTRTVNDNNECRLYKTRICEVRPCQPYSSLSLKKKSKTKKSPKPYRFTY 300
 DB 241 KTCGTGISTRTVNDNNECRLYKTRICEVRPCQPYSSLSLKKKSKTKKSPKPYRFTY 300
 QY 301 AGCLSVKXKTRPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
 DB 301 AGCLSVKXKTRPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIOSCKYVNC 360
 QY 361 HANEAAFPYRLFNIDHKFRD 381
 DB 361 HANEAAFPYRLFNIDHKFRD 381

RESULT 6
 ID AAB43987 standard; Protein; 455 AA.

XX AAB43987;

XX 08-FEB-2001 (first entry)

XX Human cancer associated protein sequence SEQ ID NO:1432.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytosolic; proliferative; vulnenerary; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antitachytic; antilyral;

KM antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KM dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KM vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KM immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KM neurological disease; drug screening.
 OS Homo sapiens.
 PN WO200055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PE 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX WPI: 2000-587533/55.
 DR N-PSDB; AAC78196.
 XX
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 11; Page 2116-2118; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 CC in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnerable; immunomodulator;
 CC antidiabetic; antiallergic; antirheumatic; antithyroid;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.
 XX
 SO Sequence 455 AA;
 Query Match 100.0%; Score 2115; DB 21; Length 455;
 Best Local Similarity 100.0%; Pred. No. 6.3e-154;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 KTCGTGISTRTYNTNPECRLYKERTICVPRCGQPVYSSLLKKKKCKSKTKSPSPVRETY 300
 Db 315 KTCGTGISTRTYNTNPECRLYKERTICVPRCGQPVYSSLLKKKKCKSKTKSPSPVRETY 374
 QY 301 AGCSLVKKRYRKYCGSCVDGRCTPOLRTYKMFRCDEGTFFSKNMVMIQSCKNYCP 360
 Db 375 AGCLSVKKRYRKYCGSCVDGRCTPOLRTYKMFRCDEGTFFSKNMVMIQSCKNYCP 434
 QY 361 HANEAPFPYRLFNDIHKFRD 381
 Db 435 HANEAPFPYRLFNDIHKFRD 455
 RESULT 7
 AAM35730
 ID AAM35730 standard; Protein; 381 AA.
 AC
 XX AAM35730;
 AC
 XX 27-MAR-1998 (first entry)
 DT
 XX Human cysteine rich protein 61 (Cyrl61).
 DE
 XX
 KW Cysteine rich protein 61; Cyrl61; human;
 KW extracellular matrix signalling molecule; cell adhesion;
 KW cell migration; cell proliferation; angiogenesis; chondrogenesis;
 KW oncogenesis; haematostasis; wound healing; organ regeneration.
 XX
 OS Homo sapiens.
 PN WO9733995-A2.
 XX
 PD 18-SEP-1997.
 XX
 PF 14-MAR-1997; 97WO-US04193.
 XX
 PR 15-MAR-1996; 96US-0013958.
 XX
 PA (MONT-) MONTIN CORP.
 PI Lau LF;
 XX WPI: 1997-470875/43.
 DR N-PSDB; AAT94699.
 XX
 PT Isolated and purified cysteine rich protein 61, Cyrl61 - useful to
 PT modulate e.g. haematostasis, induce wound healing, promote organ
 PT regeneration etc
 XX
 PS Claim 2; Page 112-113; 133pp; English.
 XX
 CC This protein sequence comprises human cysteine rich protein 61
 CC (Cyrl61), an extracellular matrix signalling molecule. Its amino
 CC acid sequence was deduced from a human placental cDNA clone (see
 CC AAT94699). Cyrl61 polypeptides can be expressed in transformed or
 CC transfected host cells. Cyrl61 can be used to modulate
 CC haematostasis, induce wound healing in a tissue, promote organ
 CC regeneration, improve tissue grafting or promote bone or prothesis
 CC implantation (claimed). It can also be used to screen for a
 CC modulator of angiogenesis, chondrogenesis, oncogenesis, cell
 CC adhesion, cell migration, cell proliferation, expand a population
 CC of undifferentiated haematopoietic stem cells in culture and to
 CC screen for a mitogen (claimed). Ex vivo methods for using
 CC mammalian extracellular matrix signalling molecules to prepare
 CC blood products are also provided.
 XX
 SO Sequence 381 AA;
 Query Match 99.6%; Score 2106; DB 18; Length 381;
 Best Local Similarity 99.5%; Pred. No. 2.5e-153;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLKAPGVGLVRDGGCCGCKVCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLKAPGVGLVRDGGCCGCKVCAKOL 60
 QY 61 NEDCSKTOPCDHRTKGLKCNFNGASSTALKGICRAQSGRCENSRITYONGESFOPKCHQ 120
 DB 61 NEDCSKTOPCDHRTKGLKCNFNGASSTALKGICRAQSGRCENSRITYONGESFOPKCHQ 120
 QY 121 CTCIDGAVGCIPLCPDELSPNLGCPNPRIVKYTGCCCEWYCDDESIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPDELSPNLGCPNPRIVKYTGCCCEWYCDDESIKDPMEDDGLLG 180
 QY 181 KELGPDASEVELTRNNELIAGVKGSSLRPLVFGMEPRILYNPLQGOKCIYVOTTSWQCS 240
 DB 181 KELGPDASEVELTRNNELIAGVKGSSLRPLVFGMEPRILYNPLQGOKCIYVOTTSWQCS 240
 QY 241 KTCGISTRTVNDNECHLVKTRICEVRPCGOPYSSILKKGKSKTKKSPKPEPRFTY 300
 DB 241 KTCGISTRTVNDNECHLVKTRICEVRPCGOPYSSILKKGKSKTKKSPKPEPRFTY 300
 QY 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFCEDGETFSKNVMMIQSCKNYVNC 360
 DB 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFCEDGETFSKNVMMIQSCKNYVNC 360
 QY 361 HANEAFPPYRLFNDIHKFRD 381
 DB 361 HANEAFPPYRLFNDIHKFRD 381

RESULT 8

AE05921
 ID AE05921 standard; Protein; 381 AA.

XX AC AE05921;

DT 24-SEP-2001 (first entry)

XX Human cysteine-rich protein (Cyr61).

XX Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KW fibroblast secreted protein; Fisp12; connective tissue growth factor;
 KW CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 KW angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KW heart disease; fibrosis; gene therapy; human.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Region 153..229

FT Domain /note="Cysteine free region"

FT Domain /note="Domain III"

FT Domain /note="Domain IV"

PN WO200155210-A2.

PD 02-AUG-2001.

PF 31-JAN-2001; 2001WO-US03267.

PR 31-JAN-2000; 2000US-0495448.

PR 15-MAY-2000; 2000US-0204364.

PR 06-OCT-2000; 2000US-0238705.

XX (MUNT-) MUNTIN CORP.

XX Lau LF, Yeung C, Greenspan JA;

XX WPI; 2001-465561/50.

DR N-PSDB; AAD1121.

XX Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods

PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration
 PS Claim 30; Page 171-172; 186pp; English.

CC The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (Fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin
 CC receptor interaction. Modulator of Cyr61-integrin alpha3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis; aberrant granulation tissue development; aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is human Cyr61
 CC protein.

SQ Sequence 381 AA;

Query Match 99.6%; Score 2106; DB 22; Length 381;
 Best Local Similarity 99.5%; Pred. No. 2.5e-153;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLKAPGVGLVRDGGCCGCKVCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRALSTCPAACHCPLKAPGVGLVRDGGCCGCKVCAKOL 60
 QY 61 NEDCSKTOPCDHRTKGLKCNFNGASSTALKGICRAQSGRCENSRITYONGESFOPKCHQ 120
 DB 61 NEDCSKTOPCDHRTKGLKCNFNGASSTALKGICRAQSGRCENSRITYONGESFOPKCHQ 120
 QY 121 CTCIDGAVGCIPLCPDELSPNLGCPNPRIVKYTGCCCEWYCDDESIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPDELSPNLGCPNPRIVKYTGCCCEWYCDDESIKDPMEDDGLLG 180
 QY 181 KELGPDASEVELTRNNELIAGVKGSSLRPLVFGMEPRILYNPLQGOKCIYVOTTSWQCS 240
 DB 181 KELGPDASEVELTRNNELIAGVKGSSLRPLVFGMEPRILYNPLQGOKCIYVOTTSWQCS 240
 QY 241 KTCGISTRTVNDNECHLVKTRICEVRPCGOPYSSILKKGKSKTKKSPKPEPRFTY 300
 DB 241 KTCGISTRTVNDNECHLVKTRICEVRPCGOPYSSILKKGKSKTKKSPKPEPRFTY 300
 QY 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFCEDGETFSKNVMMIQSCKNYVNC 360
 DB 301 AGCLSVKRYRKYCGSCVDGRCTPQLTRTVKMRFCEDGETFSKNVMMIQSCKNYVNC 360
 QY 361 HANEAFPPYRLFNDIHKFRD 381
 DB 361 HANEAFPPYRLFNDIHKFRD 381

RESULT 9

ABB09202
 ID ABB09202 standard; Protein; 374 AA.

XX AC ABB09202;

DT 08-JUL-2002 (first entry)

XX HCGF CNN family protein sequence SEQ ID NO:12.

XX Human; small CCN-like growth factor; SCGF; vulnerrary; osteopathic;
 KW gene therapy; muscle wasting disease; osteoporosis; wound healing;
 KW tissue regeneration; angiogenesis.

XX

OS Unidentified.
 XX
 PN US2002049304-A1.
 XX
 PD 25-APR-2002.
 XX
 PF 14-MAY-2001; 2001US-0853625.
 XX
 PR 06-JUN-1995; 95US-0468847.
 PR 01-APR-1998; 98US-0053587.
 XX
 PA (HAST/) HASTINGS G A.
 PA (ADAM/) ADAMS M D.
 XX
 PI Hastings GA, Adams MD.
 DR WPI: 2002-382150/41.
 XX
 PT Novel isolated polynucleotide sequence encoding a human small CCN-like
 PT growth factor, useful for treating muscle wasting disease, and
 PT osteoporosis -
 XX
 PS Disclosure; Fig 2A-D; 33pp; English.
 XX
 CC The present invention describes human small CCN-like growth factor
 CC (SCGF). SCGF has vulnerrary and osteopathic activities, and can be used
 CC in gene therapy. The SCGF polypeptides and polynucleotides can be used
 CC for treating muscle wasting diseases, and osteoporosis, and to stimulate
 CC wound healing and tissue regeneration, to promote angiogenesis and to
 CC stimulate proliferation of vascular smooth muscle and endothelial cell
 CC production. The present sequence represents a CNN family protein which
 CC is given in comparison with the human SCGF in the exemplification of the
 CC present invention.
 XX
 SQ Sequence 374 AA:
 Query Match 93.6%; Score 1980.5; DB 23; Length 374;
 Best Local Similarity 96.3%; Pred. No. 1e-143;
 Matches 360; Conservative 1; Mismatches 12; Indels 1; Gaps 1;

QY 1 MSSRIARALAVTLHLTRLALSTCPACACPLEAPKCAVGGLVRDGGCCCKVCARQL 60
 DB 1 MSSRIARELAVTLHLTRVGLSTCPADCHCPLKAPKAGVGLVRDGGCCCKVCARQL 60
 QY 61 NEDSKTOPCHTGTGECNFGASSSTALKGICRAOSEGRPCENSRITONGSEFQNCXHQ 120
 DB 61 NEDSKTOPCHTGTGECNFGASSSTALKGICRAOSEGRPCENSRITONGSEFQNCXHQ 120
 QY 121 CTCIDGAVG-CIPLCPQELSLPNLGCPNRLVKTGOCCEWVCDDESIKPMEDQGL 179
 DB 121 CTCIDGAVG-CIPLCPQELSLPNLGCPNRLVKTGOCCEWVCDDESIKPMEDQGL 179
 QY 121 CTCIDGAVG-CIPLCPQELSLPNLGCPNRLVKTGOCCEWVCDDESIKPMEDQGL 180
 DB 121 CTCIDGAVG-CIPLCPQELSLPNLGCPNRLVKTGOCCEWVCDDESIKPMEDQGL 180
 QY 180 GKEIGFDASEVELTRNNELIANGKSSILKRLPVFGEPRILYNPLQGKCIYQTTWSQC 239
 DB 181 GKEIGFDASEVELTRNNELIANGKSSILKRLPVFGEPRILYNPLQGKCIYQTTWSQC 240
 QY 240 SKTGCTGTSTRTVDNPECRIVKTRICEVRPCGOPYSSILKKGCKTKKSPRPVFT 299
 DB 241 SKTGCTGTSTRTVDNPECRIVKTRICEVRPCGOPYSSILKKGCKTKKSPRPVFT 300
 QY 300 YAGLSVYKRYPKYKCGSCVDGRCTPQLTFTVKNNRFGCEDEFTSKNMMIOSCKCNYNC 359
 DB 301 YAGLSVYKRYPKYKCGSCVDGRCTPQLTFTVKNNRFGCEDEFTSKNMMIOSCKCNYNC 360
 QY 360 PHANEAAPFYRLF 373
 DB 361 PHANEAAPFYRLF 374

RESULT 10
 AAR90919
 ID AAR90919 standard; Protein: 375 AA.
 XX

AC AAR90919;
 XX
 DT 25-JUN-1996 (first entry)
 XX
 DE Connective tissue growth factor-2.
 XX
 KW CTGF-2; connective tissue growth factor-2; secreted protein;
 KW cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
 KW insulin-like growth factor; fibroblast growth factor; Crg61.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..24
 FT /label= signal_peptide
 FT Protein 25..375
 FT /label= mature_protein
 XX
 PN W09601896-A.
 XX
 PD 25-JAN-1996.
 XX
 PF 12-JUL-1994; 94WO-US07736.
 XX
 PR 12-JUL-1994; 94WO-US07736.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Adams MD, LI H;
 XX
 DR WPI: 1996-097626/10.
 DR N-PSDB: AAT12653.
 XX
 PT Connective tissue growth factor-2 and DNA encoding it - useful to
 PT enhance the repair of connective and support tissue, and to enhance
 PT wound healing
 XX
 PS Claim 1; Fig 1A-C; 46pp; English.
 XX
 CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
 CC (AAT12653) isolated from a human foetal lung cDNA library. The CTGF
 CC polypeptides are structurally and functionally related to a family
 CC of growth factors which include IGF (insulin-like growth factor),
 CC PDGF (platelet-derived growth factor), and EGF (fibroblast growth
 CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
 CC to Crg61. Crg61 is a growth factor-inducible immediate early gene
 CC initially identified in serum-stimulated mouse fibroblasts. It encodes
 CC a member of an emerging family of secreted proteins which are also a
 CC group of cysteine-rich proteins. This group of GPs are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth.
 XX
 SQ Sequence 375 AA:
 Query Match 93.6%; Score 1980.5; DB 17; Length 375;
 Best Local Similarity 96.3%; Pred. No. 1e-143;
 Matches 360; Conservative 1; Mismatches 12; Indels 1; Gaps 1;

QY 1 MSSRIARALAVTLHLTRLALSTCPACACPLEAPKCAVGGLVRDGGCCCKVCARQL 60
 DB 1 MSSRIARELAVTLHLTRVGLSTCPADCHCPLKAPKAGVGLVRDGGCCCKVCARQL 60
 QY 61 NEDSKTOPCHTGTGECNFGASSSTALKGICRAOSEGRPCENSRITONGSEFQNCXHQ 120
 DB 61 NEDSKTOPCHTGTGECNFGASSSTALKGICRAOSEGRPCENSRITONGSEFQNCXHQ 120
 QY 121 CTCIDGAVG-CIPLCPQELSLPNLGCPNRLVKTGOCCEWVCDDESIKPMEDQGL 179
 DB 121 CTCIDGAVG-CIPLCPQELSLPNLGCPNRLVKTGOCCEWVCDDESIKPMEDQGL 180
 QY 180 GKEIGFDASEVELTRNNELIANGKSSILKRLPVFGEPRILYNPLQGKCIYQTTWSQC 239
 DB 181 GKEIGFDASEVELTRNNELIANGKSSILKRLPVFGEPRILYNPLQGKCIYQTTWSQC 240

```

QY      240 SKTCGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKSKTKKSPPEVRF 299
      |||
Db      241 SKTCGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKSKTKKSPPEVRF 300
QY      300 YAGCLSVKRRPYKPCGSCVDGRCTPQLRTVYKMRRCCEGEFTFSKNVMIIQSCCKNYNC 359
      |||
Db      301 YAGCLSVKRRPYKPCGSCVDGRCTPQLRTVYKMRRCCEGEFTFSKNVMIIQSCCKNYNC 360
QY      360 PHANEAPFPYRLF 373
      |||
Db      361 PHANEAPFPYRLF 374

RESULT 11
AAV31620
ID      AAV31620 standard; Protein; 375 AA.
XX
AC      AAV31620;
XX
DT      02-NOV-1999 (first entry)
XX
DE      Human CTGF-2.
XX
KM      Connective tissue growth factor-2; CTGF-2; wound healing; bone disorder;
XX      skin disorder; acne; burn; UV damage; stabilisation; tissue implant.
XX
OS      Homo sapiens.
XX
FH      Key
FT      Peptide      1..24      Location/Qualifiers
FT      Protein      25..375
FT      Misc-difference 268
FT      /note="Cys encoded by ICR"
XX
PN      US945300-A.
XX
XX      31-AUG-1999.
XX
PF      02-JUN-1995; 95US-0459101.
XX
PR      02-JUN-1995; 95US-0459101.
PR      12-JUL-1994; 94MO-US07736.
XX
XX      (ADAM/) ADAMS M D.
XX      (LIH/) LI H.
XX
PI      Adams MD, LI H;
XX
DR      WPI, 1999-508171/42.
XX      N-PSDB; AA211720.
XX
PT      Polynucleotides encoding growth factor polypeptides useful for
PT      enhancing the repair of connective tissue and support tissue
XX
XX      Claim 1; Fig 1; 20pp; English.
XX
XX      This sequence represents human connective tissue growth factor-2
XX      (CTGF-2). CTGF-2 cDNA was isolated from a cDNA library derived from
XX      human foetal lung. In one instance, the cDNA was cloned into a
XX      baculovirus expression vector, having first been amplified and modified
XX      via PCR using primers AA211721 and AA211722. In another instance, the
XX      cDNA was cloned into a COS cell expression vector, with prior
XX      amplification and modification using PCR primers AA211723 and AA211724.
XX      CTGF-2 is structurally and functionally related to a family of growth
XX      factors which include IGF (insulin-like growth factor), PDGF
XX      (platelet-derived growth factor) and FGF (fibroblast growth factor). This
XX      emerging family of cysteine-rich secreted proteins are important for
XX      normal growth, differentiation, morphogenesis of the cartilaginous
XX      skeleton of an embryo and cell growth. Their functions also include wound
XX      healing, tissue repair, implant fixation and stimulating increased bone
XX      mass. CTGF-2 may be used to enhance the repair of connective tissue and

```

```

CC      support tissue and can therefore treat skin disorders e.g., acne, aging,
CC      UV damage or burns. CTGF-2 can be used to promote the attachment,
CC      fixation and stabilisation of tissue implants inserted during
CC      reconstructive surgery, and can be used to enhance the healing of
CC      external wounds. It can be used in the treatment of injured or depleted
CC      bone as it promotes the growth of connective tissue, bone and cementum
CC      and stimulates protein and collagen synthesis.
XX
SQ      Sequence      375 AA;
XX
Query Match      93.6%; Score 1980.5; DB 20; Length 375;
Best Local Similarity 96.3%; Pred. No. 1e-143;
Matches 360; Conservative 1; Mismatches 12; Indels 1; Gaps 1;
QY      1 MSSRIARALAVYTLHLTRIALSTCPACCHCPLEAPKCAPGVGLYRDGCGCCYCAKOL 60
      |||
Db      1 MSSRIARALAVYTLHLTRVGLSTCPACHCHLEAPKAPGVGLYRDGCGCCYCAKOL 60
QY      61 NEDCSKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRCEYNSRIYQNGESFQPNCKHQ 120
      |||
Db      61 NEDCRKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRCEYNSRIYQNGESFQPNCKHQ 120
QY      121 CTCIDGAVG-CIPLCQDELSTPLGCPNPRIVYVVGCCCEWYCDSDSIKDPWEDDGLL 179
      |||
Db      121 CTCIGWRKACIPLCQDELSTPLGCPNPRIVYVVGCCCEWYCDSDSIKDPWEDDGLL 180
QY      180 GKELGFDAEVELTRNNELIYAVGSSSLKRLPYFGNEPRLIYPLDGGCKIYQTTSMSC 239
      |||
Db      181 GKGLGFDAEVELTRNNELIYAVGSSSLKRLPYFGNEPRLIYPLDGGCKIYQTTSMSC 240
QY      240 SKTCGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKSKTKKSPPEVRF 299
      |||
Db      241 SKTCGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCKSKTKKSPPEVRF 300
QY      300 YAGCLSVKRRPYKPCGSCVDGRCTPQLRTVYKMRRCCEGEFTFSKNVMIIQSCCKNYNC 359
      |||
Db      301 YAGCLSVKRRPYKPCGSCVDGRCTPQLRTVYKMRRCCEGEFTFSKNVMIIQSCCKNYNC 360
QY      360 PHANEAPFPYRLF 373
      |||
Db      361 PHANEAPFPYRLF 374

RESULT 12
AAE18108
ID      AAE18108 standard; Protein; 375 AA.
XX
AC      AAE18108;
XX
DT      07-MAY-2002 (first entry)
XX
DE      Human alternative connective tissue growth factor-2 (CTGF-2).
XX
XX      Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
XX      ischaemia; restenosis; tissue repair; wound healing; congenital defect;
XX      cardiovascular disease; atherosclerosis; heart failure; angina; trauma;
XX      burns; osteoporosis; periodontal disease; liver failure; tranquillizer;
XX      vulnery; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
XX      gene therapy.
XX
OS      Homo sapiens.
XX
PN      WO200204480-A2.
XX
PD      17-JAN-2002.
XX
PF      11-JUL-2001; 2001MO-US21799.
XX
PR      11-JUL-2000; 2000US-217402P.
PR      18-MAY-2001; 2001US-291642P.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX      (TRGE) TRANSGENE SA.

```


D _b	179	--LGLDSEVELTRNNELIAIGKGSILKRLPVGFTEPRVLENPFLAHGOKITVGTSMQ	236
Q _Y	239	CSMTGCGTSTRTNTPCECLVYKTRICEVAPCGQPIYSSILKGGKCKSKTKKSPDEVRF	298
D _b	237	CSMSCGGSTRTNTNPCLRLKTRICEVAPCGQPIYSSILKGGKCKSKTKKSPDEVRF	299
Q _Y	299	TYAGCSLVKKYRPRKYGSCVDRGCTPOLTRTVKKRFRCEDEGETFSKNVMIQSCKCNYN	358
D _b	297	TYAGCSLVKKYRPRKYGSCVDRGCTPOLTRTVKKRFRCEDEGETFSKNVMIQSCKCNYN	356
Q _Y	359	CPHANFAPEPYELFENDIHKFRD	381
D _b	357	CPHPNBSRFLYSLENDIHKFRD	379

ID	AAE05920	standard; Protein; 379 AA.
XX	AAE05920;	
XX	24-SEP-2001	(first entry)
XX	Mouse cysteine-rich protein (Cyr61).	
XX	Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;	
XX	fibroblast secreted protein; Fisp12; connective tissue growth factor;	
XX	CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation	
XX	angiogenesis; wound healing; Integrin receptor; atherosclerosis; tumour	
XX	heart disease; fibrosis; gene therapy; mouse.	
XX	Mus musculus.	
XX	Key	Location/Qualifiers
XX	Region	164..226
XX	FT	/note="Cysteine free region"
XX	FT	224..240
XX	Domain	/note="Domain III"
XX	FT	
XX	FT	
XX	MO200155210-A2.	
XX	02-AUG-2001.	
XX	31-JAN-2001; 2001WO-US03267.	
XX	31-JAN-2000; 2000US-0485448.	
XX	15-MAY-2000; 2000US-0204364.	
XX	06-OCT-2000; 2000US-0238705.	
XX	{MUNI-} MUNIN CORP.	
XX	Lau LF, Yeung C, Greenspan JA;	
XX	WPI; 2001-465561/50.	
XX	DR	N-PSDB; AAD11220.
XX	Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods	
XX	PT	for screening for modulators of cell adhesion, fibroblast cell
XX	PT	proliferation, angiogenesis and cell migration
XX	Example 1; Fig 1; 186pp; English.	
XX	The invention relates to extracellular matrix (ECM) signalling	
XX	molecules involved in cellular response to growth factors. More	
XX	particularly the invention is directed to cysteine-rich protein	
XX	(Cyr61), and Cyr61-related proteins such as fibroblast secreted	
XX	protein (Fisp12) and connective tissue growth factor (CTGF) and	
XX	CC nucleic acid molecules encoding such proteins. The polypeptides	
XX	of the invention are members of cysteine-rich secreted protein	
XX	family. Human Cyr61 fragment is useful in methods for screening	
XX	modulators of cell adhesion, cell migration, fibroblast cell	
XX	proliferation, angiogenesis, wound healing and Cyr61-integrin	

CC	receptor interaction. Modulator of Cyr61-integrin alphaVbeta3
CC	interaction is used for the preparation of a medicament for the
CC	treatment of atherosclerosis, heart disease, tumour metastasis,
CC	fibrosis, tumour growth, disorders associated with inadequate
CC	angiogenesis; aberrant granulation tissue development; aberrant
CC	fibroblast growth and wounds. Polynucleotides of the invention
CC	are useful in gene therapy. The present sequence is mouse Cyr61
CC	protein.
xx	
SQ	Sequence 379 AA;
	Query Match 91.6%; Score 1938; DB 22; Length 379;
	Best Local Similarity 91.4%; Pred. No. 1.8e-140;
	Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2

Db	1	MSSTFRLIAVAATVLLHLITRIALSTCPACHCPLEAPFCAGAVGLVDRGGCCCKYCAQOL	60
Qy	61	NEDCSKTQPCDHTKGLBECNFGASSTALKGICRAOSEGRPCENNSHLYONGESFQPNCKHQ	120
Db	61	NEDCSKTQPCDHTKGLBECNFGASSTALKGICRAOSEGRPCENNSHLYONGESFQPNCKHQ	120
Qy	121	CTCIDGAVGCTPLCPQELSLPLNLGCPNRLKYVYGQCCSEEWVCDEDSIKDPMEDODGLG	180
Db	121	CTCIDGAVGCTPLCPQELSLPLNLGCPNRLKYVYGQCCSEEWVCDEDSIKDPMEDODGLG	180
Qy	181	KEISGPDASEVELTRNNNELIAVGKSSSLKRLPVFGMEPRILYNPL--OGCKIVQTTSMQ	238
Db	179	--LGIDASEVELTRNNNELIAVGKSSSLKRLPVFGMEPRILYNPL--OGCKIVQTTSMQ	236
Qy	239	CSKTCGTGISTRTVNDNPECHLVKETRICEVRRPCGQPYSSLLKGGKCSKTKKSPPEVRF	298
Db	237	CSKSCGTGISTRTVNDNPECHLVKETRICEVRRPCGQPYSSLLKGGKCSKTKKSPPEVRF	296
Qy	299	TYASCLSVKKYRPXYCGSCVDGRCTPOLTRTYKMARFCEDETFPSKVMYMIOSCKNYN	358
Db	297	TYACSSVKKYRPXYCGSCVDGRCTPLQTRTYKMARFCEDETFPSKVMYMIOSCKNYN	356
Qy	359	CPHANEAAFPYRLFNDIHKFRD 381	
Db	357	CPHNEASFRLYSLFNDIHKFRD 379	

RESULT 15

ABB09201

ID ABB09201 standard; protein: 379 AA.

XX ABB09201;

DT 08-JUL-2002 (first entry)

XX Mouse cyr6 CNN family protein sequence SEQ ID NO:11.

DE Human; small CCN-like growth factor; SCGF; vulnery; osteopathic;

KW gene therapy; muscle wasting disease; osteoporosis; wound healing;

KW tissue regeneration; angiogenesis.

XX Mus sp.

OS US2002049304-A1.

XX

PN 25-APR-2002.

PD

XX 14-MAY-2001; 2001US-0853625.

PF

XX 06-JUN-1995; 95US-0468847.

PR

XX 01-APR-1998; 98US-0053587.

XX

PA (HAST/) HASTINGS G A.

PA (ADAM/) ADAMS M D.

XX Hastings GA, Adams MD;

XX WPI: 2002-382150/41.

DR Novel isolated polynucleotide sequence encoding a human small CCN-like
XX growth factor, useful for treating muscle wasting disease, and
PT osteoporosis -
PT

PS Disclosure: Fig 2A-D; 33pp; English.

XX
CC The present invention describes human small CCN-like growth factor
CC (SCGF). SCGF has vulnary and osteopathic activities, and can be used
CC in gene therapy. The SCGF polypeptides and polynucleotides can be used
CC for treating muscle wasting diseases, and osteoporosis, and to stimulate
CC wound healing and tissue regeneration, to promote angiogenesis and to
CC stimulate proliferation of vascular smooth muscle and endothelial cell
CC production. The present sequence represents a CCN family protein which
CC is given in comparison with the human SCGF in the exemplification of the
CC present invention.

XX Sequence 379 AA:

SO Query Match

91.6%; Score 1938; DB 23; Length 379;

Best Local Similarity 91.4%; Pred. No. 1.8e-140;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSRIARALAVYTLHLTRALSTCPACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSSFRTLVAVYTLHLTRALSTCPACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDSKTQPCDHTKGLCECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGSEFQPNCKHQ 120
DB 61 NEDSKTQPCDHTKGLCECNFGASSTALKGICRAQSEGRPCPEYNSRIYONGSEFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVYKVTGCCCEWYCDDESIKDPMDQDGLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIAGCPNRLVYKVTGCCCEWYCDDESIKDPMDQDGLG 180
QY 181 KELGFDAEVEELTRNNELIVAGKSSSLKRLPVFGMEPRILYNP--OGOKCIYOTTSMSQ 238
DB 179 --LGLDASEVELTRNNELIVAGKSSSLKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMSQ 236
QY 239 CSKTCGTGISTRVNDNPECRVLYETRICEVRCGQPVYSSLKKGKCSKTKKSPEVPRF 298
DB 237 CSKSCGTGISTRVNDNPECRVLYETRICEVRCGQPVYSSLKKGKCSKTKKSPEVPRF 296
QY 299 TYAGCLSVKKRYKRYCGSCYDGRCTPQLTRTVKMRRCDEGETFSKNVMMIOSCKCNYN 358
DB 297 TYAGCSSVKKRYKRYCGSCYDGRCTPQLTRTVKMRRCDEGETFSKNVMMIOSCKCNYN 356
QY 359 CPHANEAAPPEYRLFNDIHKFRD 381
DB 357 CPHPNEASFRILYSLFNDIHKFRD 379

Search completed: July 8, 2003, 12:28:53
Job time : 72 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	1938	91.6	379	2	A35669	gene CYR61 protein
2	1699	80.3	375	2	A41428	CEF-10 protein protein
3	956.5	45.2	348	2	A40578	beta IG-M2 protein
4	951	45.0	349	2	A40578	connective tissue
5	855.5	40.4	351	2	S20078	NOV protein - child
6	834.5	39.5	357	2	I38069	gene NOVH protein
7	175.5	8.3	1620	2	T27283	hypothetical protein
8	174	8.2	1111	2	T26972	hypothetical protein
9	159	7.5	1700	2	S08167	Balblab ring 3 p
10	156	7.4	1574	2	T113954	MEGF6 protein - rat
11	153	7.2	837	2	A42112	mucln-like peptid
12	152.5	7.2	1178	1	A39804	thrombospondin pr
13	152	7.2	5376	2	T42215	zonadhesin - mous
14	147.5	7.0	1170	2	A53612	lamilin Blk chain
15	145	6.9	601	2	T22025	hypothetical prot
16	145	6.8	601	2	D89711	protein F40E10.4
17	144	6.9	1964	2	T09059	notchM - mouse
18	142	6.7	1025	2	T42626	secreted leucine-r
19	141.5	6.7	1034	2	UC3598	mucln - rat
20	141.5	6.7	1101	2	T16840	hypothetical prot
21	141.5	6.7	13288	2	T03099	mucln, submucilla
22	141	6.7	3020	2	A43932	mucln 2 precursor
23	140.5	6.6	2703	1	A24430	notch protein - fi
24	139	6.6	1042	2	A55734	mucln 5AC (clone)
25	138	6.5	251	2	A55035	cysteine-rich pro
26	138	6.5	2437	2	S42612	transmembrane pr
27	136	6.4	2531	2	S18168	notch protein hom
28	136	6.4	4135	2	T42629	tenascin-X - bovin
29	135	6.4	473	2	A56175	adhesin-X - bovin

30	135	6.4	1168	2	T56985	kalinin B1 - mouse
31	134	6.3	1056	2	A53767	mucin MUC5B, trache
32	134	6.3	3002	2	A47221	fibrillin 1 precurs
33	133.5	6.3	1106	2	T13938	gene shuttle crafte
34	133.5	6.3	1480	2	A36665	slit protein 1 prete
35	133.5	6.3	1955	1	AGCH	agrin precursor -
36	133	6.3	1170	2	A40558	thrombospondin 1 p
37	133	6.3	1296	2	T16859	hypothetical protei
38	133	6.3	2813	1	VWHD	von Willebrand fac
39	133	6.3	3106	1	S53868	laminin alpha-2 ch
40	133	6.3	3672	2	T23433	hypothetical prote
41	133	6.3	3704	2	T37316	probable laminin a
42	132.5	6.3	1106	2	T44598	hypothetical prote
43	132.5	6.3	1847	2	T18308	probable vitelloge
44	132.5	6.3	2555	2	A40043	notch protein homom
45	132	6.2	2823	2	T23064	hypothetical prote

ALIGNMENTS

```

RESULT 1
A35669
gene CTR61 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 05-Nov-1999
C:Accession: A35669; I48319; S16446
R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.
Mol. Cell. Biol. 10, 3569-3577, 1990
A>Title: Expression of crr61, a growth factor-inducible immediate-early gene.
A:Reference number: A35669; MUID:90287146; PMID:2355916
A:Accession: A35669
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-379 <ONB>
A:Cross-references: GB:M32490; NID:g192909; PIDN:AAA37512.1; PID:g309206
A>Note: the authors translated the codon GAR for residue 337 as Gln
R:Iatinkic, B.V.; O'Brien, T.P.; Lau, L.F.
Nucleic Acids Res. 19, 3261-3267, 1991
A>Title: Promoter function and structure of the growth factor-inducible immediate ear
A:Reference number: I48319; MUID:91286203; PMID:2062642
A:Accession: I48319
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-379 <RES>
A:Cross-references: EMBL:X56790; NID:g50632; PIDN:CAA40109.1; PID:g50633
A>Note: the authors did not translate the codon for residue 108
A>Note: the authors translated the codon GAR for residue 337 as Gln
C:Genetics:
A:Gene: CTR61
A:Introns: 21/3; 93/1; 208/1; 279/3
C:Superfamily: von Willebrand factor type C repeat homology
F:99-166/Domain: von Willebrand factor type C repeat homology <WCV>

Query Match          91.6%; Score 1938; DB 2; Length 379;
Best Local Similarity 91.4%; Pred. No. 4.5e-135;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

Oy      1 MSRIARALAVTLLHLTRLALSTCPACACHPLEAPKCAPGVGLVDGCGCCVCAKOL 60
        |||   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1 MSSFTFRILAAVAITLLHLTRLALSTCPACACHPLEAKCAPGVGLVDGCCCKVC AKOL 60

Oy      61 NEDSKTOPCDHPTGELDECNFGASSTALKIGICRAOSEGRPCPEYNSRYONGSEFOPNCKHQ 120
        |||   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       61 NEDSKTOPCHRTGELDECNFGASSTALKIGICRAOSEGRPCPEYNSRYONGSEFOPNCKHQ 120

Oy      121 CTCIDGAVGCIPLCPOELSLPNLGCPNPRLVKYTGGCCCEBWCDSDSIKPMDQDGLG 180
        |||   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       121 CTCIDGAVGCIPLCPOELSLPNLGCPNPRLVKYSGCCCEBWCDSDSIKPSLDODDL-- 178

Oy      181 KELGFDASEVLTATNNELLAVNGSSSLKRLPVEGMEIRIIYNPL--OGOKCIYOTTSMSQ 238
        |||   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       179 --LEIDHSVEVELTATNNELLAVNGSSSLKRLPVEFTEPRVLNPLHAHGKCIVOTYSMSQ 236

```

QY	239	CSATCGTGTG	STRATNDN	NPCCRLV	KTRIC	EVN	PCG	OPY	SS	SLK	GK	CK	ST	K	SP	EV	YR	298
Db	237	CSASC	G	T	G	T	STR	T	N	D	N	P	E	C	R	L	V	296
QY	299	TYAGC	L	SV	K	K	Y	R	K	K	Y	G	C	G	C	V	D	358
Db	297	TYAGC	S	V	K	K	Y	R	K	K	Y	G	C	G	C	V	D	356
QY	359	CPHANE	A	P	F	T	R	L	N	D	I	K	E	R	D			381
Db	357	CPHANE	A	S	F	R	L	S	L	E	N	D	I	K	E	R	D	379

```

RESULT 2
A41428
CEF-10 protein precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
C:Accession: A41428
R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989
A:Title: Identification of a phorbol ester-repressible v-src-inducible gene
A:Reference number: A41428, MUID:89145206; PMID:2537491
A:Accession: A41428
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-375 <SIM>
A:Cross-references: GB:J04496; NID:g211435; PIDN:AAA4661.1; PID:g211436

```

Query Match	80.3%;	Score 1699;	DB 2;	Length 375;
Best Local Similarity	81.2%;	Pred. No. 1.6e-117;		
Matches 312;	Conservative 19;	Mismatches 41;	Indels 12;	Gaps 7

[illegible]

RESULT 3
A40578
beta IG-M2 protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
C:Accession: A40578; A53228
R:Brummer, A.; Chinn, J.; Neubauer, M.; Purchio, A.F.
DNA Cell Biol. 10, 293-300, 1991
A:Title: Identification of a gene family regulated by transforming growth factor-beta
A:Reference number: A40578; MUID:Y129699; PMID:2029337

A:Accession: A40578
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-348 <BRU>
A:Cross-references: GB:M0263; NID:g201945; PIDN:AA73135.1; PID:g201946
R:Rysek, R.P.; MacDonald-Bravo, H.; Mattei, M.G.; Bravo, R.
Cell Growth Differ. 2, 225-233, 1991
A:Title: Structure, mapping, and expression of flisp-12, a growth factor-inducible gene
A:Reference number: A53228; MUID:91363290; PMID:1888698
A:Accession: A53228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-160, 'K', 162-348 <RYS>
A:Cross-references: GB:M70641; NID:g193313; PIDN:AAA37627.1; PID:g193314
C:Genetics:
A:Gene: flisp-12

Query Match	45.28; Score 956.5; DB 2; Length 348
-------------	--------------------------------------

Matches 179; Conservative 56; Mismatches 106; Indels 41; Gaps 8;

1 MSSRIARALAVTLLHL-TRLAL-STCPACHCPL-APKCAPGVGLVRDGGGCKVCA 57

b 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCQCAEAAPHCPAGVSLVLDGGCCRVCA 60

58 KQLNEDCSKTQPCDHTKGLECNFGASSTALKGICRAQSEGRPC EYNSRIYONGESFQNC 117

61 KÖLGELECTERDPCDPHKGLFCDEGSPANRKIGVCTAK-DGAPCVFEGGSVYRSGESFÖSSC 119

118 KHQCTCIDGAVGCIPLCPQELSLPNLGCNPRLVKVTGGCCCEEWCDEDSIKDPMEDQDG 177

120 KYQCTCLDGAVGCVPLCSMDVRLPSPDCPFPRRVKLPGKCCEEWVCDEP----- 168

178 LIGKEIGFDASEVELTRNNELIavgkssLkRL-PVFGMEPRILYNPLQgQKCIvQTsw 236

169 -----KDR TAVGPALAA YRLED TFGPDPTM-----RANCLVQTTEW 205

237 S Q C S K T C G T I S T R V T N D N P E C R L V K E T R I C E V R P C G Q P V Y S S L K K G K K S K T K K S P E V 296

206 SACSKTCGMGISTRVTNDNTFCRLKQSRCLCMVRPCSEADLEENIKKKKKCIRTPKIAKPV 265

297 RFTYAGCLSVKKYRPRKYCGSCVDGRCTPQLTRTVKMRFRCEDETFSKNMMIQSCRCN 356

266 KEELSGCISVKTYRAKFCGVCTDGRCCTPHRTTLLPVEFKCPDGEIMKRNMFIKTCACH 325

357 YNCPHANEAFPEY--RLENDI 376

326 YNCPGDNDIFESLYRKMYGDM 347

RESULT A

connective tissue growth factor - human

```
;species: Homo sapiens (man)
Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text change 21-Jul-2000
```

```
;Accession: A40551; S44205
Bradham, P. M.: Tsurashi, A.: Potter, R. L.: Grotendorst, G. B.
```

. Cell Biol. 114, 1285-1294, 1991

Accession: A40551
; Reference number: A40551; MUID:913/3462; PMID:1654338

;Molecule type: mRNA
Residues: 1-349

Demar: B.S.; Werner: A.; Yang: Z.; Garner: J.M.; Gentz: B.; Innescher: T.F.; Cross-References: GB:M92934; GB:M36965; GB:S56201; NID:g180923; PIDN:AAA912/5

Description: Differential cloning and expression of human connective tissue

;Reference number: S44205
Accession: S44205

```

;Status: preliminary
Molecule type: mRNA

```

```
;Residues: 1-349 <OEM>
Cross-references: EMBL:Y78947: NID:G47A933: PTD:G47A93A
```


[illegible]

```

RESULT 5
S20078
NOV protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S20078
R:Joliot, V.; Marinette, C.; Dambine, G.; Plasiart, G.; Brisac, M.; Crochet, J.; Pert
Mol. Cell. Biol. 12, 10-21, 1992
A:Title: Proviral rearrangements and overexpression of a new cellular gene (nov) in myel
A:Reference number: S20078; MUID:92107157; PMID:1309586
A:Accession: S20078
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-351 <JOL>
A:Cross-references: EMBL:X59284; NID:g63702; PIDN:CAA41975.1; PID:g63703
C:Genetics:
A:Gene: NOV

Query Match      40.4%; Score 855.5; DB 2; Length 351;
Best Local Similarity 44.2%; Pred. No. 1.3e-55;
Matches 165; Conservative 143; Mismatches 104; Indels 61; Gaps 8;

11 LVVTLHLHTRLALS---TCPAC--HCPLPAPKCAPVGLVRGGCGGCKYCAQALNEDC 64
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
12 LLLLLLLLPRPEVSGREACPRPGGRCRPAHPRCABVPAYLDGCGCLYCAQRGESC 71
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
65 SKTPGCDHTKGLGECNFGASSTALGICRAOSEGRPCENSRIRYONGESFOPNCRKQCTCI 124
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
72 SPLPDCDESGGLYCDRGPEDGAGICMV-LEGNVCVFDGMITVRNGETFPQSKYQCTCR 130
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
125 DGAVGCIPLCQOELSLPLGLGPNRPLKAVYTGQCCSEEWDCEDDSIKDPREDQDGLIG---- 180
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
131 DGQIGCLPFCMIGILLPPDCFPFRKIEVPPECECKWVC-----DPRDEV--LLGGRFM 182
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
181 -----KELGDASEVELTRNNELLVAGKGSLSRLPYFGMEPRILYPLPGQGCIVQTT 234
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
183 AAYRQEARLGLDIVSD-----SSANCIEDTT 207
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
235 SMSQCKTCGTGISTRVTNDPNPECRLYKETRICEVRPCGQDPVYSLKKKGRKCSYTKRSP 294
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

Db	208	EMSAICSHSCGMGSTVYTRINRNOQCEKAVKOTRLCMARPEENEPSP-KKGKCIOTKSSMK	266
Qy	295	PVRFYTAGCLSVKRYKRYKSCVDRCCTPOLTRFYKMRPCDEGDETFESSNVMMIOSK	354
Db	267	AVREYFNKNCSTVQYKPRYCGLCNDRCCTPHNTITQVEFRCPOGKFLKRPMLINTCV	326
Qy	355	CNYNCPHANEAF	367
Db	327	CHGNCPOSNAFF	339

RESULT 6
138069
gene novh protein - human
C.Species: Homo sapiens (man)
C.Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C.Accession: I38069
R.Matlinierle, C.; Huff, V.; Joubert, I.; Badziach, M.; Saunders, G.; Strong, L.; Perth
Oncogene 9, 2729-2732, 1994
A.Title: Structural analysis of the human nov proto-oncogene and expression in Wilms
A.Reference number: I38069; MUID:94336229; PMID:7520150
A.Accession: I38069
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-357 <RES>
A.Cross-references: EMBL:X78351; NID:q587422; PIDN:CAAS5146.1; PID:g825696
C.Genetics:
A:Gene: novh
A:Introns: 28/3; 104/1; 188/1; 259/3
C:Superfamily: thrombospondin type 1 repeat homology
F:203-250/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match	Similarity	Score	DB 2:	Length
Best Local	100%	44.08	Pred. No. 4.7e-54;	
Matches	160;	Conservative	48;	Mismatches 117; Indels 39; Gaps 8
Qy	7	RALATVLTLLHLTLALST--CPAC--HCPLEAPKCAVGLVRDGGCCVKCAKOLNE	62	
Db	14	QCLCTFLILLHLGGVAATQRCPCPPCPRGCRPATPTPCAFGVAVALDGGSCCLVCARQRE	73	
Qy	63	DCSKTPQPCDHTKGLGCNFGASSTALKGICRAOSEGRPCPEYSRIYONGESFQPNCKHOCT	122	
Db	74	SCSDLEPCDESSGLICDRSADPSNQTGICTA-VEDGNCVFDGVITRSRSGKFPSPCKFOCT	132	
Qy	123	CIDGAVGCIPLCPQELSLPNIAGCPNPRLVKVTGQCCCEMVCNEDSIKDPMEDQDGLLKE	182	
Db	133	CRDQIGICVPRCOLVILLPEPNCAPARKVEYGECECKWICGPD-----EDSLGLLT	185	
Qy	163	IGFDSAEVELTNNNELIAGKSSKLRLPFQMERILYNPLQGGKCIQYTTSSQSCST	242	
Db	186	IA--AYRREPATLGVSV---SDSV-----NCIEQTTEMTWCAKSKS	219	
Qy	243	CGTGISTRTNDNPCECLYKEDRIEVRPCGQ-PYVSSLKKGKCKSKTKSPEDPRFYTA	301	
Db	220	CGMGSTRTNNRQCEEMKQTRILDMVRCEDEPDPDPTDKKCKCLRTKSKLKAITHLQF	279	
Qy	302	GCLSYKKRPYKYGSCVDRCCTPOLTRIVYKRRFRCEDEFTSKNVMYIOSCKANCYCPH	361	
Db	280	NOTSLHTYKPRRCGVCSDRCCCTPINTKTIQAFQCSPOQIYKKRPVMYIGTCTCHTNCPK	339	
Qy	362	ANEA 365		
Db	340	ANEA 343		
RESULT 7				
T27283				
hypothetical protein Y64GI0A.f - Caenorhabditis elegans				
C:Species: Caenorhabditis elegans				
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999				
C:Accession: T27283				
R:Almscough, R.				

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20336

A:Accession: T27283

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1620 <WIL>

A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CAB54471.1; CESP:y64G10A.f

A:Experimental source: clone Y64G10A

C:Genetics:

A:Gene: CESP:y64G10A.f

A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1; 7

Query Match

Best Local Similarity 8.3%; Score 175.5; DB 2; Length 1620;
Matches 92; Conservative 33; Mismatches 146; Indels 109; Gaps 21;

```

QY 26 CPACACPLEAP-----KCAPG-----VGLVRDCCGCGCKYCAKQALNEDCK 66
DB 1052 CKKTCGQNGATCDSTVSGCECRPGMRGKKCDRPPCDG--RREGCNALICDCTTTVDTSN 1109
QY 67 TOP---CDHTKGLCECNFGASSTALGICRAOSEGRCEVNSRIYONGSEFQNCXHOCT 122
DB 1110 YNFVARCHDHVWG-EGR-----CPAGWTGPDQCTSCPLGRHGE---GCRHSQ 1153
QY 123 CIDGAVGCIPLCPQELSLNRLGCPNRLVYVGCCEBNVCDSDSIKDPMEDODGLGKE 182
DB 1154 CSGNA-----SCDRVTGF--CDPSPGMRKNCSECEPGLMGSNCKHCLCMHSGCNKE 1206
QY 183 LGFDASEVELTRNNELIAGVGSLSLRLPVFGMEPRILYNPLQGKCIYOT-----TSM 236
DB 1207 NG-----DCEIDIDMTGPSLCPPOFGFRNCARCNCKNAGASCDKRTGRCGLPGW 1256
QY 237 S--QCSKTGCTGISTRTVNDNPECLVKEFRICEVPRCGQPVYSSLK-----KXKCSKT 289
DB 1257 SEHCHEKSCVSG-----HYGAKC---EECECENALCDPISGHSCQPGMRGKKCNRP 1307
QY 290 KKSPEVRYTYAGCLSVKRYRYCG-----SC--VDGACCTPOLTRYKMAFRCE 338
DB 1308 -----CL--KGYFGHCHSOSGRCANSSKCDHISGRCCPKGYAGHSCTELCP 1352
QY 339 DGETFSKNVMMIQSCCKNYN 358
DB 1353 DG-TFGESES--QKCDGGEN 1369

```

RESULT 8

T26972

hypothetical protein Y47H9C.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T26972

R:Haris, B.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26972

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1111 <WIL>

A:Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GM00019; CESP:y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:y47H9C.4

A:Map position: 1

A:Introns: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match

Best Local Similarity 8.2%; Score 174; DB 2; Length 1111;
Matches 87; Conservative 30; Mismatches 164; Indels 142; Gaps 20;

```

QY 26 CPAACHPLE-APKAPGVGLVR-----DGGCGCKYCAK-QALNEDCKTPDCHTKGLEC 78
DB 380 GSKTCTVARENTLMCAPINTGFCRCKPGFYGDNCCLASKDSYGPNCERQAMCDMWHASPC 439

```

```

QY 79 NFASSTALKGICRAOSEGRPEVNSRIYONGSEFQPNCKHOCTC-----IDGA----- 127
DB 440 NNETGSC-----VCKPRTGKNCSBPCL-----DFYGPNCAMHQCQGNQRGVGDGADGKQ 491
QY 128 -----VGLIPL-----CPQELSLPN--LIGCP-- 146
DB 492 CDGWTGHNRCHEHPCADPTFGANCKRCKPKGIGCDPRTGECTCPAGLOGANCIDGCPBG 551
QY 147 -----NPLVAVTGQCCCEBNVCDSDSIKDPMEDODGLGKEFGDASEVELT 193
DB 552 SYGPGCKLKCKYNGKCDKETGEC-----TC-----QPFPGSPCSTTCSKRGK 596
QY 194 RNNELIAGVGSLSLRLPVFGMEPRILYNPLQGKCIYOT--TSMGCSKTGCTG--IS 248
DB 597 ESCCELSPCSDASCSK-----QTKCLCPGLTKGVSCDQKCDPNTFGFLC 641
QY 249 TRVTDNPECLVKEFRICEVPRCGQPVYSSLSKKGKCKTKKSPRPVFTYAGCLSVK 308
DB 642 QETVTPSPCASPDPKNGVCLSCPBG--SGIHCHNC-----PAGSYDGCQY-- 688
QY 309 YRPKYGSCVDGRCCTPOLTRY-----KMRFCPEDETFESKNVMMIQSCCKNYNCP 360
DB 689 -----C-SCADHGCDPTTGBCICEPRTGKTCSEKCPGKTYGICALCPKASBSTD 742
QY 361 HAN 363
DB 743 HIN 745

```

RESULT 9

S08167

Balbiant ring 3 protein - midge (Chironomus tentans)

C:Species: Chironomus tentans

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000

C:Accession: S08167

R:Faulstich, G.; Lendahl, U.; Gallit, J.; Ericsson, C.; Wieslander, L.

A:Title: The balbiant ring 3 gene in Chironomus tentans has a diverged repetitive str

A:Reference number: S08167; MID:90172404; PMID:1689777

A:Accession: S08167

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-1700 <PAU>

A:Cross-references: GB:X52263; NID:97057; PIDN:CAA36506.1; PID:97058

C:Genetics:

A:Gene: BR3

A:Map position: 4

C:Superfamily: unassigned Balbiant ring proteins

Query Match

Best Local Similarity 7.5%; Score 159; DB 2; Length 1700;
Matches 90; Conservative 46; Mismatches 169; Indels 116; Gaps 22;

```

QY 30 CHCPLPAK--CAPGVGLVRCGCG-CKYCAKQALNEDCKTPDCHTKGLEC-----NFG 81
DB 221 CICIPAEPAAGCSAPLKMDDKCSGACPAKMEKKEKESKINPNPTCEGCAQLNCP 280
QY 82 AASTALKGICRAO--SEGRCEVNSRIYONGSEFQPNCKHOCTC----- 124
DB 281 DNKKRANKETQCECKEVKRC-----NGQYF--CKDSCSCVCPGCGDKRTCTAPQYV 330
QY 125 DGAAGCIPLCPELSPNIGCPNRLVYVGTGQCCBNVCDSDSIKDPMEDODGLGKEIG 184
DB 331 DG-VACSCSGPYNMOKPADGCRP-----QKMDREBRRCBPVK-HCKRNGKYMWD 378
QY 185 FDASEVELTRNNELIAGVG--GSS-----IKRLVFGMEPRILYNPLQGKCIYOTTSW 236
DB 379 ETICQICPRDAPVCTAGKERGCEGCKCINREKEGCAPLVWNE--NTCKVCVPADK- 436
QY 237 SGCSTGCTGISTRTVNDNPECLVKEFRICEVPRCGQPVYSSLSKKGKCKTKKSPRPVFTYAGCLSVK 280
DB 437 QMSPGCGSGKSFNKLTCQCECDGASAKGILKRNADYCKCECQPGMPEPGGKQYTWISD 496

```

QY 281 KKQKSKT-----KSPERFAYAGCLSVKKYRP---311
 Db 497 KKCCESSPTTCQAPOLILNICECKCPVNMALQKECKSPKMTDCKCECSTPATC 556
 QY 312 ---KYCG---SCV---DGRCTPOLTRTKMRFCDEDET-FSKNVMMIQSCKCNVNC 359
 Db 557 ESKQTMCGEACQICICGGDKNCGNKKFEDKPSCECKKNPTCTSPQVMADNCEC--KC 614
 QY 360 P 360
 Db 615 P 615

RESULT 10

MEGF6 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T13954
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
 A:Reference number: Z14126; MUID:98360089; PMID:9693030
 A:Accession: T13954
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 11574 <NAK>
 A:Cross-references: EMBL:AB011532; NID:93449293; PIDN:BA32462.1; PID:93449294
 A:Experimental source: strain Sprague-Dawley; brain
 C:Genetics:
 A:Gene: MEGF6

Query Match 7.48; Score 156; DB 2; Length 1574;
 Best Local Similarity 21.98; Pred. No. 0.0014;
 Matches 97; Conservative 37; Mismatches 132; Indels 176; Gaps 28;

QY 26 CPACHCP-----LEAPKAPGV--GLVRDQ-----50
 Db 568 CSSPCTCQNGCTDPPVLAGRCPCPGVSGANCDGCKPFGKCKCKCHCANRCRRLY 627
 QY 51 GCKKCAKOL-----NEDCSKTQPCDHTKGLKCNFGASSTALKICR 92
 Db 628 GAC-LCPGLYGRFCHLACPPMAFGPCSEDCLEQ--SHTRSCNPKDSCS-----CK 678
 QY 93 AAGEGRCEYNSRIYQNGESFOENCKHOCICIDGAVCIPL-----CPQ 136
 Db 679 AGFQGERQAE--ESG-FFGPCRRCTCQPG-VACDPVSGECRTQCPPGYQGEDCGQ 733
 QY 137 ELSLPNIG-----CPNRLKVTGQC-CEEWVCDDESIKPMEDODGILGELGFA 187
 Db 734 ECVVGTGVCNCSGSCSVGAPCHRYTQECICPCKGTGEDGADCPERRMGLGCEI-----789
 QY 188 SEVELTRNNELIANGKSSIKRLPVFGMEPRILYNLOGKCIYQTTSMQSKNTC---243
 Db 790 -----CPACEHGASC-----NPEGTGCLCLPGFVSGRCDQTCAGW 825
 QY 244 -GTGISTR--VTNDNPECRILVETRICVEVPCQPVYSSLKKGCKSKTKKSPVAVFTY 300
 Db 826 YGTGCOIRACANDG-HCDPTTGRCSC-----APGWTGLISCFACDSGHWDPD-----872
 QY 301 AGCLSVKYPKYC---GSC--VDGR-CT-----POLRTVYMRP---RCEDEGFYSK 345
 Db 873 --CI-----HPCNSAGHGNCDVNSGLCLCEAGIEGRCDSCROGTYYGSCF-----918
 QY 346 NVMMIOSCKCNYN--CPHANE 365
 Db 919 -----QKRCHEHGAACDHVSGA 935

RESULT 11
 A42112
 mucin-like peptide MLP 2677 - rat

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Oct-1997
 C:Accession: A42112
 R:Xu, G.; Huan, L.J.; Khatri, I.A.; Wang, D.; Bennick, A.; Fahm, R.E.; Forstner, G.G.
 J. Biol. Chem. 267, 5401-5407, 1992
 A:Title: cDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide
 A:Reference number: A42112; MUID:92184794; PMID:1371999
 A:Accession: A42112
 A:Status: preliminary
 A:Molecule type: nucleic acid; protein
 A:Residues: 1-837 <XU1>
 A:Experimental source: intestine
 A:Note: sequence inconsistent with the nucleotide translation
 C:Superfamily: von Willebrand factor type C repeat homology
 F:582-650/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.28; Score 153; DB 2; Length 837;
 Best Local Similarity 19.48; Pred. No. 0.0012;
 Matches 89; Conservative 39; Mismatches 138; Indels 192; Gaps 24;

QY 25 TCQA-----ACHCPLEAPKAPGVGLVRDQCG-----52
 Db 419 TCFAHROYACGPESEPTCQSSPKNSTLLVEGCFCEGTTKRAPGYDVCVKICGCVGP 478
 QY 53 -----CKVC-----AKQINDCSKTQPCDH 72
 Db 479 NVPRERGEHFPDCKDVCLEGGSGIVCOPKRCARGLTTCEDGTLYLVLEADPDKCN 538
 QY 73 TKGLECN-----FGASSTALKIGICRAOSEGRP---CEYNSRIYONGESFOP 115
 Db 539 TTSCOKDPRCKKAKRPSCLLGFVEKSEHVPGKCPYSCVPKGVCHENMEYDGPSPVYS 598
 QY 116 NCKHQCICD-----GAVGC--IPLCQELSLPMLGCPNP-RIVKVTGCCCEWVCD 164
 Db 599 NKCDQCVCTSMDSNSTOLNVICTHP-C-----NISCSGFEIYEVGECCKQQT 650
 QY 165 EDGIKDPMD-----QDGLLCKELG-----PDASVELTRNNELIANGKSSILRLPVFGM 215
 Db 651 HCLTKRQDQYITLKEIQRNPNDRCTFFSKMT---NNQL-----SSVNIITCPDF 701
 QY 216 EPRILYNLOGKCIYQTTSM--SQCSKTGCTGISTRTVNDNPECRILVETRICVEVPCG 273
 Db 702 DP-----SPCVPSITLYMPNNGCCKTC-----IHPN-----NTVPCS 733
 QY 274 Q-PYSSLKKGKCKSKTKKSPVRYRTYACLSYKKRKYC-GSC-----VD 319
 Db 734 AIPVMEI-----SYNGC--AKNISMFCAGSCGTFAVMSAQAQDLD 773
 QY 320 G--RCCTPOLTRVYMRFCDEDETFSKNVMMIQSCK 355
 Db 774 HGSCCNEERTSVAMVSLDPCDGSKLSHTHTIESCLC 811

RESULT 12
 A39804
 thrombospondin precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A39804
 R:Lawler, J.; Duquette, M.; Ferro, P.
 J. Biol. Chem. 266, 8039-8043, 1991
 A:Title: Cloning and sequencing of chicken thrombospondin.
 A:Reference number: A39804; MUID:91217026; PMID:2022631
 A:Accession: A39804
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1178 <LAW>
 A:Cross-references: GB:M60853; NID:9212763; PIDN:AAA51437.1; PID:9212764
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
 F:325-383/Domain: von Willebrand factor type C repeat homology <VWC>
 F:386-437/Domain: thrombospondin type 1 repeat homology <THRI>
 F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>

hypothetical protein F40E10.4 - *Caenorhabditis elegans* (fragment)

C:Species: *Caenorhabditis elegans*

C:date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T22025

R:Stuye, R.

submitted to the EMBL Data Library, February 1996

A:Reference number: 219503

A:Accession: T22025

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-601 <MIL>

A:Cross-references: EMBL:Z69792; PIDN:CAA93668.1; GSPDB:GN00028; CESP:F40E10.4

A:Experimental source: clone F40E10

C:Genetics:

A:Gene: CESP:F40E10.4

A:Map position: x

Query Match

6.9%; Score 145; DB 2; Length 601;

Best Local Similarity 19.8%; Pred. No. 0.0034;

Matches 93; Conservative 52; Mismatches 142; Indels 182; Gaps 24;

```
QY 30 CHCLELPKCAPGVGLVRDGGC---CKVCAKOLNEDCSKTQPC----- 70
    |||:|
Db 162 CDCPME-----YEGHCEDKLEYCTKRLN-PCENNGKCIPINGSYSCMCSPGFT 209
    |||:|
QY 71 -----DHTKGLECNFGASSTALKGI-----CRAQSEGRPCENYSRI---YONGESF 113
    |||:|
Db 210 GNNCETNIDCKNVECCNGGS--CYDGLSTDCRCRPGYAGQCYETIPPMADMEYQKTDAC 267
    |||:|
QY 114 QPNCCKHO-----CTCIDGAVGCIPLCPQELSLPILGCPNP----- 148
    |||:|
Db 268 QQSACGGGECVVASQNSSDFTCKCHEGFSG--PSCDRQMS---VGFKNPGAYLALDPLASD 322
    |||:|
QY 149 -----RLYKVTGQCCEWVCDEDSIKDPMED 174
    |||:|
Db 323 GTITMPLRTYSGIGILLYGDDHFFVSAELYDGRVKLYIIGNFPASHMYSVKVNDGLPH 382
    |||:|
QY 175 QDGLLGRK---LGFDASEVELTRN---NELIAYGK----- 203
    |||:|
Db 383 RISTRTSERKCFLOIDKNPQIVVNSGKSDOLITKGKEMLYIGGLPIEKSODAKRRPHVK 442
    |||:|
QY 204 -GSSLKRLPVFGMEPRILYN--PLQGQKCIYQVTTWSQCSKT-----CGTGISTRVT 252
    |||:|
Db 443 NSESLK-----GCISITINEVPIINQDALENVNTQESATVNFCAIDCGNG--KCT 494
    |||:|
QY 253 ND--NPECRLVKEIRICEVRPCGQPVYSSLKKGKCSKTKSPPEVPTVAGCLSVKKYR 310
    |||:|
Db 495 NNALSPKGYMCQ---CDSHESGE--HCDEKRIKCDKQKFRRHIE--NECRSVDRIK 544
    |||:|
QY 311 ---PKYGCSCVDGRCTPPLTRTVYKMRFRCEDEFTFSKNYMIQSCKC 355
    |||:|
Db 545 IAEONGYCGG--EONCCIAVKKKQKRYMKICNGTKISTVHITRCCGC 591
    |||:|
```

Search completed: July 8, 2003, 12:31:40
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 8, 2003, 12:13:12 ; Search time 23 Seconds

(without alignments)
687.064 Million cell updates/sec

Title: US-09-901-910-2

Perfect score: 2115
Sequence: 1 MSSRIARALALVLTLLHLTR.....ANEAAPFYRLPNDIKHFRD 381

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES -

Result No.	Score	Query Match	Length	ID	Description
1	2115	100.0	381	1	CYR6_HUMAN
2	1938	91.6	379	1	CYR6_MOUSE
3	1699	80.3	375	1	CE10_CHICK
4	952.5	45.0	348	1	CTGF_MOUSE
5	951	45.0	349	1	CTGF_HUMAN
6	937	44.3	349	1	CTGF_BOVIN
7	910	43.0	349	1	CTGF_PIG
8	855.5	40.4	351	1	NOV_CHICK
9	852	40.3	353	1	NOV_COTJA
10	834.5	39.5	357	1	NOV_HUMAN
11	829.5	39.2	343	1	NOV_XENLA
12	823	38.9	354	1	NOV_MOUSE
13	159	7.5	1700	1	BAR3_CHITE
14	153	7.2	837	1	BAR3_CHITE
15	152.5	7.2	1178	1	TSP2_CHICK
16	152	7.2	5376	1	ZAN_MOUSE
17	151	7.1	3110	1	LMA2_HUMAN
18	148.5	7.0	2282	1	ZAN_RABIT
19	144.5	6.8	1172	1	LMB3_HUMAN
20	144	6.8	1964	1	NTC4_MOUSE
21	143	6.8	1173	1	TSP1_XENLA
22	141	6.7	5179	1	MUC2_HUMAN
23	140.5	6.6	2703	1	NOTC_DROME
24	138	6.5	453	1	HRA3_HUMAN
25	138	6.5	2437	1	NTC1_BRARE
26	138	6.5	2437	1	VWF_PIG
27	137.5	6.5	480	1	HRA1_MOUSE
28	137	6.5	1170	1	TSP1_BOVIN
29	137	6.5	1587	1	LMG3_HUMAN
30	136	6.4	2531	1	NTC1_RAT
31	136	6.4	2813	1	VWF_CANFA
32	135	6.4	473	1	FP2_MYGA
33	135	6.4	1168	1	LMB3_MOUSE

34	134	6.3	1056	1	MUC5_HUMAN	P98088 homo sapien
35	134	6.3	1218	1	JAG1_MOUSE	O99xx0 mus musculu
36	134	6.3	2871	1	FBN1_HUMAN	P35555 homo sapien
37	133.5	6.3	1106	1	STC_DROME	P40798 drosophila
38	133.5	6.3	1480	1	SLIT_DROME	P24014 drosophila
39	133.5	6.3	1955	1	AGRI_CHICK	P31696 gallus gall
40	133	6.3	460	1	HRA3_MOUSE	O9d236 mus musculu
41	133	6.3	1170	1	TSP1_MOUSE	P35441 mus musculu
42	133	6.3	2813	1	VWF_HUMAN	P04275 homo sapien
43	133	6.3	3106	1	LMA2_MOUSE	O60675 mus musculu
44	133	6.3	3672	1	LMT2_CAEEL	O21313 caenorhabd
45	131.5	6.2	703	1	FBN1_HUMAN	P23142 homo sapien

ALIGNMENTS

RESULT 1
CYR6_HUMAN STANDARD: PRT: 381 AA.
ID CYR6_HUMAN 000622: 014934: 0A3775: 09BZL7:
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61) (GIG1
DE protein) (Insulin-like growth factor-binding protein 10).
CN CYR61 OR IGFBP10 OR GIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Aldrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RX MEDLINE-97280750: Pubmed-9135077;
RA Jay P., Berge-Lefranc J.L., Marsollier C., Mejean C., Travaux S.,
RA Beta P.;
RT "The human growth factor-inducible immediate early gene, CYR61, maps
RT to chromosome 1p.";
RL Oncogene 14:1753-1757(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-98197344: Pubmed-9536281;
RA Martinerie C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
RT "Chromosomal mapping and expression of the human cyr61 gene in tumour
RT cells from the nervous system.";
RL Mol. Pathol. 50:310-316(1997).
RN [3]
RP SEQUENCE FROM N.A.
RA Kolesnikova T.V., Lau L.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Bl A.B., Yu L.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Schuetz N., Lechner A., Groll C., Koehnle J., Jakob F.;
RT "Regulation of hcyr61 by vitamin D, serum and cytokines in fetal human
RT osteoblasts.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;
RT "Organization and expression of the cyr61 gene in normal human
RT fibroblasts.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.

CC TISSUE-Lung, Placenta, and Skin;
 RA Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY. CEF-10/CYR61/CTFG/ETSP-12/NOV PROTEIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y12084; CAA72802.1; -
 DR EMBL: U62015; AAB58319.1; -
 DR EMBL: Y11307; CAA72167.1; -
 DR EMBL: AF003594; AAB61240.1; -
 DR EMBL: AF003385; AAB84227.1; -
 DR EMBL: Z98053; CAB10848.1; -
 DR EMBL: AF07860; AAG59863.1; -
 DR EMBL: BC001271; AAH01271.1; -
 DR EMBL: BC009199; AAH09199.1; -
 DR EMBL: BC016952; AAH16952.1; -
 DR Gene: HGNC:2654; CYR61.
 DR MIM: 602369; -
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00090; tsp_1.1.
 DR Pfam: PF00093; vwc_1.
 DR Pfam: PF00219; IGFBP.1.
 DR SMART: SM0041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; VWC.1.
 DR PROSITE: PS00222; IGF-BINDING.1.
 DR PROSITE: PS01185; CTCK_1.1.
 DR PROSITE: PS01225; CTCK_2.1.
 DR PROSITE: PS01208; VWF_C.1.
 DR Growth factor binding: signal.
 DR SIGNAL: 1 34
 FT CHAIN 1 34
 FT DOMAIN 25 381
 FT DOMAIN 98 164
 FT DISULFID 286 360
 FT DISULFID 303 337
 FT DISULFID 314 353
 FT DISULFID 317 355
 FT DISULFID 322 359
 FT DISULFID 165 165
 FT CONFLICT 210 210
 FT CONFLICT 220 220
 FT CONFLICT 369 369
 FT SEQUENCE 381 AA; 42026 MW; FCOBD39C078CA0B1 CRC64;
 Query Match 100.0%; Score 2115; DB 1; Length 381;
 Best Local Similarity 100.0%; Pred. No. 3.3e-150;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 61 MEDCKTQPCDHTKGLKLECNFGASSTALKICRAQSGRCEYNRIYONGESFOPNCKHQ 120
 QY 121 CTCIDAVAGCIPICPQELSLPNLGCNPNRLVKTGQCCCEWYCDDESIDPMDODGLG 180
 DB 121 CTCIDAVAGCIPICPQELSLPNLGCNPNRLVKTGQCCCEWYCDDESIDPMDODGLG 180
 QY 181 KELGPDASVELTRNNELIAGKSSSLKRLPVFGMEPRILYNPLGOKCIYOTTSWQCS 240
 DB 181 KELGPDASVELTRNNELIAGKSSSLKRLPVFGMEPRILYNPLGOKCIYOTTSWQCS 240
 QY 241 KTCGTSIRRVNDNPECLVETRICVPCGOPYSSSLKKGKSKTKSPKPEYRTY 300
 DB 241 KTCGTSIRRVNDNPECLVETRICVPCGOPYSSSLKKGKSKTKSPKPEYRTY 300
 QY 301 ACCLSVKRYRPGYSGVDCRCCTPOLTRTVKMRFCDEGETFSKNVMAIQSCKNYCP 360
 DB 301 ACCLSVKRYRPGYSGVDCRCCTPOLTRTVKMRFCDEGETFSKNVMAIQSCKNYCP 360
 QY 361 HANEAPFPYRLPNDIKFRD 381
 DB 361 HANEAPFPYRLPNDIKFRD 381
 RESULT 2
 ID CYR6_MOUSE STANDARD; PRT; 379 AA.
 AC P18406;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (insulin-like growth factor-binding protein 10) (3CH61).
 GN CYR61 OR IGFBP10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_Taxid=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/C; TISSUE-Fibroblast;
 RX MEDLINE-90287146; PubMed-2355916;
 RA O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
 RT "Expression of cyr61, a growth factor-inducible immediate-early
 gene.";
 RL Mol. Cell. Biol. 10:3569-3577(1990).
 RN (12)
 RP SEQUENCE FROM N.A.
 RC STRAIN-AJ; TISSUE-embryonic fibroblast;
 RX MEDLINE-91288203; PubMed-2062642;
 RA Latick B.V., O'Brien T.P., Lau L.F.;
 RT "Promoter function and structure of the growth factor-inducible
 immediate early gene cyr61.";
 RL Nucleic Acids Res. 19:3261-3267(1991).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
 CC -1- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
 CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
 CC IN LUNG.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
 CC NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
 CC -1- INDUCTION: BY GROWTH FACTORS.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY. CEF-10/CYR61/CTFG/ETSP-12/NOV PROTEIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----


```

CC -----
DR EMBL; M32490; AAA37512.1; -
DR EMBL; X56790; CAA40109.1; -
DR PIR; A35669; A35669.
DR MGI; 88613; Cyf61.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00219; IGFBP; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWC; 1.
DR Growth factor binding; Signal.
DR SIGNAL; 1 24
DR CHAIN; 25 379
DR DOMAIN; 98 164
DR DISULFID; 284 321
DR DISULFID; 301 335
DR DISULFID; 312 351
DR DISULFID; 315 353
DR DISULFID; 320 357
SQ SEQUENCE 379 AA; 41709 MW; FA6B5014B5A8EE9 CRC64;

```

```

Query Match 91.6%; Score 1938; DB 1; Length 379;
Best Local Similarity 91.4%; Pred. No. 4,5e-137;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

```

```

QY 1 MSSRIARALAVLTLHLTRLATSTCPACACPLAEPKAPGVGLVYDGGCCCKVCAKOL 60
DB 1 MSSSTFRLAVAVTLHLTRLATSTCPACACPLAEPKAPGVGLVYDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFPQNCKH 120
DB 61 NEDCSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFPQNCKH 120
QY 121 CTCTIDGAVGCIPLCPQELSLPNLGCNPNRLVKYTGCCCEEWVCDSDIKDMEDODGLG 180
DB 121 CTCTIDGAVGCIPLCPQELSLPNLGCNPNRLVKYTGCCCEEWVCDSDIKDMEDODGLG 180
QY 181 KELGFDASEVELTERNNELIAGKSSSLKRLPFVFGMEPRILVNP--OGOKCIYOTTSMSQ 238
DB 179 --LGLDASEVELTERNNELIAGKSSSLKRLPFVFGTEPVVLNPLHAHQKCIYOTTSMSQ 236
QY 239 CSKTCTGISTRTVNDNPECRVLETRICEVRPCGOPYSSLSKKGKCSKTKKSEPVRF 298
DB 237 CSKCGGISTRTVNDNPECRVLETRICEVRPCGOPYSSLSKKGKCSKTKKSEPVRF 296
QY 299 TYAACLSTVKKRYKRCSCVNGRCCTPOLITVYKMRFCGEGLFSPKVMVMIQSCKNYN 358
DB 297 TYAGCSSVKKRYKRCSCVNGRCCTPOLITVYKMRFCGEGLFSPKVMVMIQSCKNYN 356
QY 359 CPHANEAAPFYRLFNDIHKFRD 381
DB 357 CPHENEASFRLYSIFNDIHKFRD 379

```

```

RESULT 3
CE10.CHICK STANDARD; PRT; 375 AA.
AC P19336;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE CEF-10 protein precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89145206; PubMed=2537491;
RA Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.;
RT "Identification of a phorbol ester-repressible v-src-inducible gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
CC -1- INDUCTION: BY V-SRC.
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
CC PROTEIN FAMILY. CEF-10/CTRG/FTSP-12/NOV PROTEIN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 WFC DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; J04496; AAA48661.1; -
DR PIR; A41428; A41428.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00219; IGFBP; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWC; 1.
DR Growth factor binding; Signal.
DR SIGNAL; 1 22
DR CHAIN; 23 375
DR DOMAIN; 98 164
DR DOMAIN; 281 318
DR DISULFID; 298 332
DR DISULFID; 309 348
DR DISULFID; 312 350
DR DISULFID; 317 354
SQ SEQUENCE 375 AA; 40651 MW; 95F28553BE35D5AE CRC64;

```

```

Query Match 80.3%; Score 1699; DB 1; Length 375;
Best Local Similarity 81.2%; Pred. No. 2,5e-119;
Matches 312; Conservative 19; Mismatches 41; Indels 12; Gaps 7;

```

```

QY 1 MSSRIARALAVLTLHLTRLATSTCPACACPLAEPKAPGVGLVYDGGCCCKVCAKOL 59
DB 1 MGSACARP-ALAAALICLARALAGSPCAVOCPPAANAQCAFGVGLVYDGGCCCKVCAKOL 59
QY 60 LNEDCSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFPQNCKH 119
DB 60 LNEDCSKTQPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRIVONGESFPQNCKH 119
QY 120 QCTCTIDGAVGCIPLCPQELSLPNLGCNPNRLVKYTGCCCEEWVCDSDIKPMEDODGLL 179
DB 120 QCTCTIDGAVGCIPLCPQELSLPNLGCNPNRLVKYTGCCCEEWVCDSDIKPMEDODGLL 177

```

QY 180 GKELGPDASEVELLRNNELAVGKSSLRPFVFGMP--RLYNPILOGOKIVQTTSSMS 237
 Db 178 SEGEGLDASEGELLRRNNELLAIVYG-GLKMLPVFGSSPOSRAEPNP----KCIYQTTSSMS 232
 QY 238 QCSKTCGTGISTRTVNDNPECLVETRICVPRCGGQPVYSSLRKGGKSKTKSPSPVR 297
 Db 233 QCSKTCGTGISTRTVNDNPECLVETRICVPRCGGQPVYSSLRKGGKSKTKSPSPVR 292
 QY 298 PTAAGCLSVKRYRKYRGSCVDRCCTPOLRTVYKMFRCDEGTFESKNVMIQSCKNY 357
 Db 293 FTVAGCASSVKRYRKYRGSCVDRCCTPOLRTVYKMFRCDEGTFESKNVMIQSCKNY 352
 QY 358 NCPHANEAFPEYRLFNDIKFRD 381
 Db 353 NCPHANE-AYPEYRLFNDIKFRD 375
 RESULT 4
 CTGF_MOUSE STANDARD; PRT; 348 AA.
 AC P29266;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Connective tissue growth factor precursor (CTGF) (FISP-12 protein)
 GN CTGF OR FISP12 OR FISP-12 OR HCS24.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91363290; PubMed=1888698;
 RA Ryseck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
 RT "Structure, mapping, and expression of fisp-12, a growth factor-
 RT inducible gene encoding a secreted cysteine-rich protein.";
 RL Cell Growth Differ. 2:225-233(1991).
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91229699; PubMed=2029337;
 RA Brunner A., Chinn J., Neuberger M.G., Purchio A.F.;
 RT "Identification of a gene family regulated by transforming growth
 RT factor-beta.";
 RL DNA Cell Biol. 10:293-300(1991).
 RN 13
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=97327410; PubMed=9184077;
 RA Kireeva M.L., Lelinkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
 RA Adler A.S., Lau L.F.;
 RT "Cyr61 and fisp12 are both ECM-associated signaling molecules:
 RT activities, metabolism, and localization during development.";
 RL Exp. Cell Res. 233:63-77(1997).
 RN 14
 RP FUNCTION.
 RX MEDLINE=99182484; PubMed=10082563;
 RA Babic A.M., Chen C.C., Lau L.F.;
 RT "Fisp12/mouse connective tissue growth factor mediates endothelial
 RT cell adhesion and migration through integrin alphavbeta3, promotes
 RT endothelial cell survival, and induces angiogenesis in vivo.";
 RL Mol. Cell. Biol. 19:2958-2966(1999).
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY).
 CC -1- FUNCTION: MEDIATES CELL ADHESION AND ENHANCES FIBROBLAST GROWTH
 CC FACTOR-INDUCED DNA SYNTHESIS. PROMOTES ENDOTHELIAL CELL SURVIVAL.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
 CC SOLUBLE FORM.
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 CC -1- INDUCTION: BY GROWTH FACTORS.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING

CC PROTEIN FAMILY: CEF-10/CYR61/CTGF/FISP-12/NOV PROTEIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M70641; AAA37627.1; -;
 DR EMBL: M70642; AAA37628.1; -;
 DR EMBL: M80263; AAA73135.1; -;
 DR PIR: A53228; A53228.
 DR MGD: MGI:95537; Ctgf.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Ins1_gro_fac-pr.
 DR InterPro: IPR000884; TSPL.
 DR InterPro: IPR001007; WFC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSPL; 1.
 DR SMART: SM00214; WFC; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS01208; WFC; 1.
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 348
 FT DOMAIN 100 166
 FT DOMAIN 255 329
 FT DISULFID 255 292
 FT DISULFID 272 306
 FT DISULFID 283 322
 FT DISULFID 286 324
 FT DISULFID 291 328
 FT CONFLICT 161 161 K -> E (IN REF. 2).
 SQ SEQUENCE 348 AA; 37793 MW; 735b5b6a711686f CRC64;
 Query Match 45.0%; Score 952.5; DB 1; Length 348;
 Best Local Similarity 46.6%; Pred. No. 6.3e-64;
 Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;
 QY 1 MSSRIARALAVVTLHL-TRLAL-STCPAACHPLE-APKCAPGYGLVRDGGCCCKVGA 57
 Db 1 MLASVAGPISLALVLTALCTRPATVGODCSAQCCCAAEAPAPCHPAVSLVLDGGCCCRVCA 60
 QY 58 KQINECCSKTOPCDHDKGLGECNFGASSTALKGICRAQSGREPREYNSRIYONCESTQPMC 117
 Db 61 KQGLGELCTERDPCDPKHLGFCDFGSPANRRIGVCTAK-6GAPCVFGSVYRSSEPSQSSC 119
 QY 118 KHQCTIDGAVGICPLCPPELSLPNIGCPNRLVYTGOCCEWVDESDINDPMEDDQG 177
 Db 120 KQCTCLDPAVGVCLPSMDVRLPSDCPPRRVRLPKGCKCKRWVDEP----- 168
 QY 178 LKGLGPDASEVELLRNNELAVGKSSLRPLVFGMPRLYNPILOGOKIVQTTSSMS 236
 Db 169 -----KDRRAVGPALAAVRLIEDTFGDDPIIM-----RANGLVQTTSM 205
 QY 237 SOCSTCGTGTSTRTVNDNPECLVETRICVPRCGGQPVYSSLRKGGKSKTKSPSPVR 296
 Db 206 SACSCTCGMGISTRTVNDNPECLVETRICVPRCGGQPVYSSLRKGGKSKTKSPSPVR 265
 QY 297 RFTVAGCLSVKRYRKYRGSCVDRCCTPOLRTVYKMFRCDEGTFESKNVMIQSCKNY 356
 Db 266 KEFLSGCTSVKTYRAFCVCTDGRCTHRTTTLPEVFKCPDGLIMKNMFIKTCACH 325

DB 276 MKTYRAKFCVCTDGRCTPHRTTTLPEVFKCPDGEVKKMMNFITCACHYNCPODNDI 335

QY 366 APPFY--RLFNDI 376

DB 336 FESLYYRKMYGDM 348

RESULT 6

CTGF_BOVIN STANDARD; PRT: 349 AA.

ID CTGF_BOVIN 09GL71:

AC 018739: (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Connective tissue growth factor precursor.

GN CTGF.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RP SEQUENCE FROM N.A.

RC TISSUE=Arter:

RA Lillensiek B., Lin Z., Fotsis T., Schimanski M., Bierhaus A., Kanitz M., Kaufmann G., Schweigler L., Ziegler R., Nawroth P.P.; Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver:

RA Mathias M., Schwitters C., Hove M., Rupp S., Erond N.E.; "Bovine connective tissue growth factor, organization of the chromosomal gene and demonstration of promoter activity." Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RL [1]

CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOCHONDRIAL SECRETED BY VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY).

CC -1- FUNCTION: MEDIATES CELL ADHESION AND ENHANCES FIBROBLAST GROWTH

CC -1- FACTOR-INDUCED DNA SYNTHESIS (BY SIMILARITY).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A SOLUBLE FORM (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN FAMILY. CEF-10/CYR61/CTGF/FISP-12/NOV PROTEIN SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF000137; AAB6596.1; -

DR EMBL: AF309555; AAG30290.1; -

DR InterPro: IPR000359; Cys_knot.

DR InterPro: IPR000867; Insl_gro_fac_dr.

DR InterPro: IPR000884; TSP1.

DR InterPro: IPR001007; WVF_C.

DR Pfam: PF00007; Cys_knot; 1.

DR Pfam: PF00090; tsp; 1.

DR Pfam: PF00093; wvc; 1.

DR Pfam: PF00219; IGRBP; 1.

DR SMART: SM00041; CT; 1.

DR PROSITE: PS00222; IGF_BINDING; 1.

DR PROSITE: PS01185; CTCK_1; 1.

DR PROSITE: PS01225; CTCK_2; 1.

DR PROSITE: PS01208; WVF; 1.

KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.

FT SIGNAL 1 26 POTENTIAL.

FT CHAIN 27 349 CONNECTIVE TISSUE GROWTH FACTOR.

FT DOMAIN 101 167 WVC.

FT DOMAIN 256 330 CTCK.

FT DISULFID 256 293 BY SIMILARITY.

FT DISULFID 273 307 BY SIMILARITY.

FT DISULFID 287 325 BY SIMILARITY.

FT DISULFID 292 329 BY SIMILARITY.

FT CONFLICT 28 28 D -> DC (IN REF. 1).

FT CONFLICT 58 58 MISSING (IN REF. 1).

FT CONFLICT 88 88 A -> T (IN REF. 1).

FT CONFLICT 103 104 CV -> YI (IN REF. 1).

FT CONFLICT 166 169 CDEP -> SRDE (IN REF. 1).

FT CONFLICT 184 184 P -> L (IN REF. 1).

FT CONFLICT 200 200 L -> Q (IN REF. 1).

FT CONFLICT 209 209 C -> Y (IN REF. 1).

FT CONFLICT 269 269 E -> Q (IN REF. 1).

FT CONFLICT 284 284 C -> F (IN REF. 1).

SO SEQUENCE 349 AA; 37924 MW; 5F0C8EB3EFBA4F99 CRC64;

Query Match 44.3%; Score 937; DB 1; Length 349;

Best Local Similarity 46.4%; Pred. No. 9e-63;

Matches 172; Conservative 61; Mismatches 100; Indels 38; Gaps 8;

QY 10 ALVYTLHLRLALS-TCPAACHPL-EAPKAPGVGLVRDGGCCCKVCAKQLEDCSKT 67

DB 12 AFVLLALCISRPASSODCSAPCCCPAGPAPRCAGSLVLDGCGCCRCVCAKQLEDCSKT 71

QY 68 OPDHDHKGLECNAGSATALMGICRAGSCEVSRIRYONGESFOPCKHOCCTDGA 127

DB 72 DPDPDHKGLPCDGSFANRRIIGCTAK-DGAPVFGGVYSGESFOSCKTQCTLDGS 130

QY 128 VGCIPLCPOELSLPNIACPNRLVYVVGCCCEWDCEDSIRDPMDQDGLGKEIGPDA 187

DB 131 VGVPLVCSDVRLPSPDCPFPRVRLPGKCCCEWDC-----PKHT--VVGPAI---- 179

QY 188 SEVELTRNNELAVGSSSLKRLPVGMERPLIYNLQCKIVOTSSQSKTCGTGI 247

DB 180 -----AAVREDTFGPPTMI-----RANLVOTETWSASCKTCGMGI 217

QY 248 STRVTNDNDECRVKEIRICEVRPCQGPVYSILKKGKCKSKSPRYFTYAGLSVK 307

DB 218 STRVTNDNAFCRLKESKSRICMVRPCADLEINIKKCKIRIPKISKPIKFLSGCTSK 277

QY 308 KYRPKYSGCVDRCTPOLTRFKMRFCDEGETFSKVMYMQSCCKYNCNPANAAF 367

DB 278 TYRAKCGVCTDGRCTPHRTTTLPEVFKCPDGEVKKMMNFITCACHYNCPODNDI 337

QY 368 PFY--RLFNDI 376

DB 338 SLYYRKMYGDM 348

RESULT 7

CTGF_PIG STANDARD; PRT: 349 AA.

ID CTGF_PIG 019113:

AC 019113:

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Connective tissue growth factor precursor.

GN CTGF.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.

OX NCBI_TaxID=9823;

RP SEQUENCE FROM N.A.

RC TISSUE=Uterus;

RX MEDLINE=97390475; PubMed=9242708;

RA Bristol D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R., Harding P.A.; "Purification and characterization of novel heparin-binding growth factors in uterine secretory fluids. Identification as heparin-

RT regulated Mr 10,000 forms of connective tissue growth factor.*;
 RL J. Biol. Chem. 272:20275-20282(1997).
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOGENIC FACTOR SECRETED BY
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY).
 CC -1- FUNCTION: MEDIATES CELL ADHESION AND ENHANCES FIBROBLAST GROWTH
 CC FACTOR-INDUCED DNA SYNTHESIS (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
 CC SOLUBLE FORM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY. CEF-10/CYR61/CTGF/ETSP-12/NOV PROTEIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WWC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U83916; AAC48756.1; -.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WWC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; wwc; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; WWC; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS01208; WWC_1; 1.
 KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 349
 FT DOMAIN 101 167 WWC.
 FT DOMAIN 256 330 CTCK.
 FT DISULFID 256 293 BY SIMILARITY.
 FT DISULFID 273 307 BY SIMILARITY.
 FT DISULFID 284 323 BY SIMILARITY.
 FT DISULFID 287 325 BY SIMILARITY.
 FT DISULFID 292 329 BY SIMILARITY.
 SQ SEQUENCE 349 AA; 38007 MM; BB510E2B2B52DA40 CRC64;
 Query Match 43.0%; Score 910; DB 1; Length 349;
 Best Local Similarity 45.2%; Pred. No. 9.1e-61;
 Matches 168; Conservative 56; Mismatches 108; Indels 40; Gaps 7;
 QY 10 ALVYTLHLRLAL-STCPAACHPL-EAPKCAVGGLYRGCCGCCCAVCAQLWEDSKT 67
 DB 12 AFVLLALCSRPAGGQDCSGCQCAAGKRRACPAAGVSLVYDGCRCRCAQLDELCTER 71
 QY 68 QPCDHTKLECNFASSTALKGICRAOSEGRPEYNSRIYONGSFQPNCHHOCTIDGA 127
 DB 72 DPCDPHGLFCDPESPAKRIGVCTAK-DGAPCFVGGIVYRSGSFSSCKYQCTCLDGA 130
 QY 128 VGCIPICPDELSPNLGCPNRLVYVKGCCCEWVCDSDSTKDPEDDGLLGELEFEDA 187
 DB 131 VGVPLCSMDVRLSPDPCPFRRVYKLPKCKCEEWVCDP----- 169
 QY 188 SEVELTRNNELIANGKSSSLKRL-PVGRMERILLYNPLOGCKIVOTTSNCSKTCGTG 246
 DB 170 -----KDHTVVGPALAAVRLDTEGDPPTMM-----RANCLVOTTEMCSKTCGGMG 216
 QY 247 ISTRVTDNDFECRLVKEIRICEVAPCGQPVYSLKKGKCKSKTKSPVAFVFTAGCLSV 306

DB 217 ISTRVTDNDFECRLVKEIRICEVAPCGQPVYSLKKGKCKSKTKSPVAFVFTAGCLSV 276
 QY 307 KTRPKYCGSCVDRCTPOLTRTVKRFCEDETFESKNVMIQSCNCPHANEAA 366
 DB 277 KTYRACGCVCTDRCTCPHRTTLTPVEFKCPDDEVAKKSMFICTACAHNCPGDNDIF 336
 QY 367 PFYF-RFLPNDI 376
 DB 337 ESTLYRMYGDM 348
 RESULT 8
 ID NOV_CHICK STANDARD; PRT; 351 AA.
 AC P28686;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
 GN NOV.
 OS Gallus gallus (chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Brown Leghorn;
 RX MEDLINE=92107157; PubMed=1309586;
 RA Joliet V., Martinie C., Dambrine G., Plassiat G., Brissac M.,
 RA Crochet J., Perbal B.;
 RT "Proviral rearrangements and overexpression of a new cellular gene
 RT (nov) in myeloblastosis-associated virus type 1-induced
 RT nephroblastomas.";
 RL Mol. Cell. Biol. 12:10-21(1992).
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
 CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION
 CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL, IN
 CC MUSCLE AND INTESTINE, IN THE EMBRYO. LUNG AND LESS SO IN BRAIN AND
 CC SPLEEN, IN ADULT CHICKEN.
 CC -1- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 CC ADULT KIDNEY.
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY. CEF-10/CYR61/CTGF/ETSP-12/NOV PROTEIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WWC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X59284; CAA41975.1; -.
 DR PIR: S20078; S20078.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WWC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; wwc; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.

DR SMART: SM00214; VWC: 1.
 DR PROSITE: PS00222; IGF_BINDING: 1.
 DR PROSITE: PS01185; CTCK: 1.
 DR PROSITE: PS01225; CTCK: 2; 1.
 DR PROSITE: PS01208; VWC: 1.
 KM Proto-oncogene; Growth factor binding; Signal.
 FT SIGNAL: 1 24
 FT CHAIN: 1 24
 FT DOMAIN: 104 170
 FT DISULFID: 258 332
 FT DISULFID: 258 295
 FT DISULFID: 275 309
 FT DISULFID: 286 325
 FT DISULFID: 289 327
 FT DISULFID: 294 331
 FT CARBOHYD: 274 274
 SO SEQUENCE 351 AA; 38268 MW; 1ECB3FA3058C6797 CRC64; (POTENTIAL).
 Query Match 40.48; Score 855.5; DB 1; Length 351;
 Best Local Similarity 44.28; Pred. No. 1e-56;
 Matches 165; Conservative 43; Mismatches 104; Indels 61; Gaps 8;
 QY 11 LVVTLHLRLALST---TCPAC--HCPLEAKCAPGVGLVDRGGCCCKVCAKOLNDEC 64
 DB 12 LLLLLLLLRPRCEVSGREAPCPRCGRCPAEPRAVPAVLDGCGCCCLVCAKRGESG 71
 QY 65 SKTOPCDHTKGLCECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGESFQPNCKHOCCTCI 124
 DB 72 SPLLPDESGLGCDKGPEDGGGAGICMV-LEBDNCFVFDGMIRNGETFQPSCKYCTCR 130
 QY 125 DGAVGCIPLCPQELSLPNLGCPRRLVRYTGGCCCEWVCDSDIKPMEDQDLG---- 180
 DB 131 DQIGICLPKCNLGLLPDPCCPRRIEYVGECEKWC-----DPREY--LLGCFAM 182
 QY 181 -----KELGFDAEVELTRNNELIANGKSSLRPLVFGMEPRILLYNPLOGOKIVOT 234
 DB 183 AAYRQATATGIDVSD-----SSANCIEDYT 207
 QY 235 SMSCKSKTGCTGISTRTVNDNPECLVKEIRICEVAPRCQPYVSSLKKGKSKTKKSP 294
 DB 208 EMSACKSCSGMGSTVYTNRNQCEWVKQTRLCMRBCENEPSD-KKKGKCIQTKKSK 266
 QY 295 PVAFYTAGCLSVAKYKPKYCGSCVDGRCTCPQLTRVYKARFRCEDETSKNNMIQSK 354
 DB 267 AAFEEKNCTSVQTYPRYCGLCNDGRCTPHNTKTIOVEFRCDPKFLKPMMLINTCV 326
 QY 355 CANYNCPHANEAF 367
 DB 327 CHGNCPSNNAF 339
 RESULT 9
 NOV_COTJA STANDARD: PRT; 353 AA.
 ID NOV_COTJA
 AC P42642;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
 GN NOV
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Weiskirchen R., Bister K.;
 RL Submitted (Aug-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY. CEF-10/CYR61/CTFG/FISP-12/NOV PROTEIN SUBFAMILY.

CC -1- SIMILARITY: CONTAINS 1 VWC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U13063; AAA21128.1; ..
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insi_gro_fac_pr.
 DR InterPro: IPR000884; TSPL.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00090; Cys_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSPL; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS00222; IGF_BINDING: 1.
 DR PROSITE: PS01185; CTCK: 1; 1.
 DR PROSITE: PS01225; CTCK: 2; 1.
 DR PROSITE: PS01208; VWC: 1.
 KM Proto-oncogene; Growth factor binding; Signal.
 FT SIGNAL: 1 26
 FT CHAIN: 27 353
 FT DOMAIN: 106 172
 FT DOMAIN: 260 334
 FT DISULFID: 260 297
 FT DISULFID: 277 311
 FT DISULFID: 288 327
 FT DISULFID: 291 329
 FT DISULFID: 296 333
 FT CARBOHYD: 276 276
 SO SEQUENCE 353 AA; 38667 MW; 717D9F8533882E89 CRC64; (POTENTIAL).
 Query Match 40.38; Score 852; DB 1; Length 353;
 Best Local Similarity 43.58; Pred. No. 1.9e-56;
 Matches 165; Conservative 44; Mismatches 104; Indels 66; Gaps 8;
 QY 10 ALVVTLLHLRLALST---TCPAC--HCPLEAKCAPGVGLVDRGGCCCKVCAK 58
 DB 8 SLPVLLLLLLLRSEVNGREAPCPRCGRCPAEPRAVPAVLDGCGCCCLVCAK 67
 QY 59 QLNEDSKTOPCDHTKGLCECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGESFQPNCK 118
 DB 68 QRGESCSPLLPDESGLGCDKGPEDGGGAGICMV-LEBDNCFVFDGMIRNGETFQPSCK 126
 QY 119 HOCCTIDAVGCIPLCPQELSLPNLGCPRRLVRYTGGCCCEWVCDSDIKPMEDQDL 178
 DB 127 YQCTCRDQIGICLPKCNLGLLPDPCCPRRIEYVGECEKWC-----DPREY--L 178
 QY 179 LG-----KELGFDAEVELTRNNELIANGKSSLRPLVFGMEPRILLYNPLOGOK 228
 DB 179 LCGFMAAAYROENTGIDVSD-----SSAN 203
 QY 229 CIVOTTSWCSKTKTGISTRTVNDNPECLVKEIRICEVAPRCQPYVSSLKKGKSK 288
 DB 204 CIEQTFEMSACSRCSMGSTVYTNRNQCEWVKQTRLCMRBCENEPSD-KKKGKCI 262
 QY 289 TKSPVPVAFYTAGCLSVAKYKPKYCGSCVDGRCTCPQLTRVYKARFRCEDETSKNNM 348
 DB 263 TRKSKMAVFEYKNCSTVQTYPRYCGLCNDGRCTPHNTKTIOVEFRCDPKFLKPM 322
 QY 349 MIOSCKCANYNCPHANEAF 367
 DB 323 LINTCVCHGNCPSNNAF 341

```

RESULT 10
NOV_HUMAN
ID NOV_HUMAN STANDARD: PRT; 357 AA.
AC P46745; Q96BYS;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NOV protein homolog precursor (NovH) (Nephroblastoma overexpressed
  gene protein).
GN NOV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94336229; PubMed=7520150;
RA Mathierie C., Huff V., Joubert I., Badeloch M., Saunders G.F.,
RA Strong L.C., Perbal B.;
RT *Structural analysis of the human nov proto-oncogene and expression
  in Wilms tumor.;
RL Oncogene 9:2729-2732(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
  GROWTH REGULATION (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE
  STROMAL TYPE.
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
  PROTEIN FAMILY. CER-10/CYR61/CTGF/FTSP-12/NOV PROTEIN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch)
CC -----
DR EMBL; X78351; CAAS5146.1; -
DR EMBL; X78352; CAAS5146.1; JOINED.
DR EMBL; X78353; CAAS5146.1; JOINED.
DR EMBL; X78354; CAAS5146.1; JOINED.
DR EMBL; X96584; CAAS5146.1; -
DR EMBL; BC015028; AAH15028.1; -
DR Gene; HGNC:7885; NOV.
DR MIM; 164958; -
DR InterPro: IPR000359; Cys_knot.
DR InterPro: IPR000867; Ins1_gro_fac_pr.
DR InterPro: IPR000884; TSPL.
DR InterPro: IPR001007; VWFC.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00093; tsp_1; 1.
DR Pfam; PF00093; vwc_1.
DR Pfam; PF00219; IGFBP; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSPL; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00222; IGF_BINDING; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWFC; 1.
KM Proto-oncogene; Growth factor binding; Signal.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 357 NOV PROTEIN HOMOLOG.

```

```

FT DOMAIN 108 174 VWFC.
FT DOMAIN 264 338 CTCK.
FT DISULFID 264 301 BY SIMILARITY.
FT DISULFID 281 315 BY SIMILARITY.
FT DISULFID 292 331 BY SIMILARITY.
FT DISULFID 295 333 BY SIMILARITY.
FT DISULFID 300 337 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 280 280 N-LINKED (GLICNA. . .) (POTENTIAL).
FT CONFLICT 97 97 N->K (IN REF. 2).
SO SEQUENCE 357 AA; 39162 MW; 035D5BF4576D85B CRC64;

Query Match 39.5%; Score 834.5; DB 1; Length 357;
Best Local Similarity 44.0%; Pred. No. 3.7e-55;
Matches 160; Conservative 48; Mismatches 117; Indels 39; Gaps 8;

OY 7 RALALVVTLLHLTRLAIST--CPAAC--HCPLEAPKCAPGVGLYRGGCGCKCAKOLNE 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 14 QCLCLTFLLHLLOVAVATORCPQCPQRCFATPPTCAPGVRAVLDCSCCLVCARORGE 73
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 63 DCSKTQPDHRTKGLGECNFGASSTALKGICRANOSRGPCEVNSRYONGESFQPNCKHQT 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 74 SCSDLEPDSSGLYCDRSAPSNQGTCTA-VEGDNCVFDGVYRSGEKFQPSCKRQCT 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 123 CIDGAVGCTPLCPQELSLPNIGCPNRLKYTGCCCEMEVDEDSIKDPMEDODGLLKE 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 133 CRDQICGVPRQQLDVLIPNCPAPKRVPEVPGCECKWICGP-----EDDSIGLT 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 183 LGFDASEVELTRNNELIANGKSSILKRLPVGMEPRILYNPLQOKCIVQTSMSQSKT 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 186 LA-AVREAVLGEV-----SDSSV-----NCTEQTEWTACSKS 219
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 243 CGTGISFVTVNDNEDCRIVKTRICEVAPRQGO-PYSSLKKGKCKSKKSPDEVRTIYA 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 CGMEFSFVTRNMQCEMLKQTRLCMPVCEDEPQDPDKKGLKTRKSLKIHLOFK 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 302 GCLSVKKRPRYKSCVNDGRCCTPOLPTFTVMRRCFDEGFESKVMVMSQCKNYNCPH 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 NCTSLHTYKPRFCGVCSGRCRCPHTNTIQAERQCSGQIVKRPVWYIGTCYCHTNCPK 339
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 362 ANEA 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 NNEA 343

RESULT 11
NOV_XENLA
ID NOV_XENLA STANDARD: PRT; 343 AA.
AC P51609;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NOV protein homolog precursor (Xnov).
GN NOV.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96257227; PubMed=8666280;
RX Ying Z., King M.L.;
RT "Isolation and characterization of xnov, a Xenopus laevis ortholog of
  the chicken nov gene.";
RL Gene 171:243-248(1996).
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
  GROWTH REGULATION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
  PROTEIN FAMILY. CER-10/CYR61/CTGF/FTSP-12/NOV PROTEIN SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 VWFC DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U37063; AAB17096.1; -
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_dr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00219; IGFBR; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS01208; VWF_C; 1.
 DR PROSITE: PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR Growth factor binding; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 343
 FT DOMAIN 93 159
 FT DOMAIN 249 323
 FT DISULFID 249 286
 FT DISULFID 266 300
 FT DISULFID 277 316
 FT DISULFID 280 318
 FT DISULFID 285 322
 FT CARBOHYD 265 265
 SQ SEQUENCE 343 AA: 38070 MW: 677D0708B21365F CRC64:
 N-LINKED (GLCNAC -) (POTENTIAL).

Query Match 39.2%; Score 829.5; DB 1; Length 343;
 Best Local Similarity 42.1%; Pred. No. 8.5e-55;
 Matches 159; Conservative 43; Mismatches 113; Indels 63; Gaps 9;

QY 9 LALVYLLHLRLALSTPCAC-HCPLEAPKAPGVGLVRDGGCCCKVCAKQINCCSKT 67
 DB 5 LALCFILL-IQVVASQKCPSCDQCEPPSCAPVLLILDGCGCPVCAKQEGESC 63
 QY 68 QPCDHTKGLCEVNGASTALKICRAQSEGRCEVSR1YONGESFQPKCHOCCTIDGA 127
 DB 64 NPGQEKGLXCEPMDPRMETGTCA-LEGNSCVFDGVYRRNRESQPSCKHYCTLN 122
 QY 128 VGCIPICPOELSLPNIGCPNRLVYVYGCCCEWVDEDSINDPMEDODGLG----- 180
 DB 123 IGVPCRCNLDLLPGDCPFPRVYKVPGRCEKWCDS--KEEM-----AIGGFAMA 174
 QY 181 ---KELGPASEVELTRNNELAVGSGSLKRLPVGMERILLINLQCKATVQTTSS 237
 DB 175 REPARIGDASDPSFA-----CAIQTTSS 199
 QY 238 QCSKTGCTISTRTVNDNEPCRLVETRICVEVPC-GQPVYSLKKKCKSKTKSPRE 296
 DB 200 ACSKTGCMGVSSVYTRNRARCEMOKOIRLCMYRSCSEEGVHEVKKGKCKVARKTK 259
 QY 297 RFTYAGCLSVKRYKPYKGVCGVDRGCTPQTLRTVYMRPCEDEGETFSKNVMISQCK 356
 DB 260 HFHYKNTCTSVQPKPFQCGSGRCCTPHSTKTMHVEVPCQKRIYKRPVWISVCH 319
 QY 357 YNCP-----HANEAAP 368
 DB 320 YNCPDSSLLQVENAF 337

RESULT 12
 NOV_MOUSE

ID NOV_MOUSE STANDARD; PRT; 354 AA.
 AC 064299;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
 DE gene protein).
 GN NOV.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV, and ICR: TISSUE=Brain;
 RX MEDLINE=97131523; Pubmed=8975721;
 RA Snaith M.R., Natarajan D., Taylor L.B., Choi C.P., Martinerie C.,
 RA Perbal B., Schofield P.N., Boulter C.A.;
 RT "Genomic structure and chromosomal mapping of the mouse nov gene.";
 RL Genomics 38:425-428(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=96204003; Pubmed=8622864;
 RA Martinerie C., Chevalier G., Kauscher F.J. III, Perbal B.;
 RT "Regulation of nov by Wt1: a potential role for nov in
 RT nephrogenesis.";
 RL Oncogene 12:1479-1492(1996).
 CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING
 CC PROTEIN FAMILY. CEF-10/CYR61/CTGF/ETSP-12/NOV PROTEIN SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X97863; CAA66457.1; -
 CC EMBL: Y09257; CAA70454.1; -
 CC EMBL: X96585; CAA65404.1; -
 CC MGD: MGI:109185; Nov.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_dr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00219; IGFBR; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VWC; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS01208; VWF_C; 1.
 KW Proto-oncogene; Growth factor binding; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 354
 FT DOMAIN 102 168
 FT DOMAIN 261 335
 FT DISULFID 261 298
 FT DISULFID 278 312
 FT DISULFID 289 328
 FT DISULFID 292 330
 FT DISULFID 297 334
 FT POTENTIAL.
 FT NOV PROTEIN HOMOLOG.
 FT VWF_C.
 FT CTCK.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.
 FT BY SIMILARITY.

FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 354 AA; 38928 MW; 08E8C8C67829DE CRC64;
 Query Match 38.9%; Score 823; DB 1; Length 354;
 Best Local Similarity 44.2%; Pred. No. 2.6e-54;
 Matches 161; Conservative 48; Mismatches 119; Indels 36; Gaps 9;

QY 7 BALAVYTLH-LTRIALS-TCRPAAC--HCPLEAPKCAPGVGLVROCGGCKYCAKOLNE 62
 DB 8 KCLCIGFLPHILISVYASLRCPSPKCPISPTAPAVRSYLDSCCPVCARORGE 67
 QY 63 DCSKTOPCDHTKGLCNFGASSTALKGICRAOSEGRECEYNSRIYONGESFOPNCKHOCT 122
 DB 68 SCSEMRRCQDSSGLYCDRSADPNNGSIGCNV-PEGDNCVFDGYIYRNGEKFEPNCGYCT 126
 QY 123 CIDGAVGCIPLCPQELSLPMLGCPNRLVYKVTGOCCEWVCDSDSIKDPMEDDGLLKE 182
 DB 127 CRDGOIGCLPRCOLDVLPPGDCPAPKVAVPGECCKMTGSGDSE----QGTGTLG-G 180
 QY 183 LGFDASVEVLTFRNNELIYAVGSSSLKRLPYFGMEPRILYNPLQOKCIYOTTSMSQSKT 242
 DB 181 LAIPARPEATVGEV---SDSSI-----NCIEQTTEWSAGSKS 216
 QY 243 CGTGTSTRTVNDNPECHLYKRETRICEVPCGO--PYVSLKKGKCKSKTKSPFVPTYA 301
 DB 217 CGMGVSTRVYNNRRKOCMAVOTRLCIYRCPQDEEYTDKKGKCLPTKSKLHIOFE 276
 QY 302 GCLSVKKYRPRKYGSCVDGRCCTPOLTRTYKMRFCEDGETFSKNVMMIOGSKCNVCPH 361
 DB 277 NCTSLTYKPRFCGVCSDGRCCTPHNTKTIOVEFOCLPGEI IKRPVAVIGTCTCYSCNPQ 336
 QY 362 ANEA 365
 DB 337 NNEA 340

RESULT 13
 BAR3_CHITE
 ID BAR3_CHITE STANDARD; PRT: 1700 AA.
 AC 003376;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Balbiani ring protein 3 precursor.
 GN BR3.
 OS Chironomus tentans (Midge).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
 OC Chironomidae; Chironomidae; Chironominae; Chironomus.
 OX NCBI_TaxID=7153;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RX MEDLINE=90172404; PubMed=1689777;
 RA Paulsson G., Lendahl U., Gall J., Ericsson C., Wieslander L.;
 RT "The Balbiani ring 3 gene in Chironomus tentans has a diverged
 repetitive structure split by many introns.";
 RL J. Mol. Biol. 211:331-349(1990).
 CC -1- FUNCTION: USED BY THE LARVAE TO CONSTRUCT A SUPRAMOLECULAR
 STRUCTURE, THE LARVAL TUBE. BALBIANI RING PROTEIN 3 COULD PLAY A
 CC ROLE AS A TRANSPORT PROTEIN THAT BINDS TO OTHER PROTEINS
 CC INTRACELLULARLY AND IN THE GLAND LUMEN IN ORDER TO PREVENT THESE
 CC FROM FORMING WATER-INSOLUBLE FIBERS TOO EARLY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: SALIVARY GLAND.
 CC -1- DOMAIN: HAS 82 APPROXIMATE REPEATS OF CYS-X-CYS-X-CYS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X52263; CA36506.1; -.
 DR PIR: S08167; S08167.
 DR HSSP: P15358; 1SK2.
 DR InterPro: IPR004153; CXKC-repeat.
 DR Pfam: PF03128; CXKC; 71.
 DR Repeat: Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 1700 BALBIANI RING PROTEIN 3.
 SQ SEQUENCE 1700 AA; 186145 MW; 34202B28521B0815 CRC64;
 Query Match 7.5%; Score 159; DB 1; Length 1700;
 Best Local Similarity 21.4%; Pred. No. 0.00024;
 Matches 90; Conservative 46; Mismatches 109; Indels 116; Gaps 22;

QY 30 CHCPLEAPK--CAPVGLVNDGCGC--CKVCAKQLNEDCSKTOPCDHTKGLGC-----NFG 81
 DB 221 CICTAPAPAGCSAPLKWDDDKSCACAPAKMEKKCKVESGKIWNPNTECGGCAQLNCP 280
 QY 82 ASSTALKGICRAQ--SEGRPEYNSRIYONGESFOPNCKHOCTCI----- 124
 DB 281 DNKKANKETQCCEKEVKKC-----NGQGVF---CKDSQCYCPGSDXDKTCTADQVY 330
 QY 125 DGAVGCIPLCPQELSLPMLGCPNRLVYKVTGOCCEWVCDSDSIKDPMEDDGLLKE 184
 DB 331 DG-VACSCSCPVNNQKADGCPRP-----QKMDKECRCECPVK-HDCNKGVMYD 378
 QY 185 FDASEVELTRNNELIYAVGK---GSS-----LKRIPVGMERILYNPLQOKCIYOTTSW 236
 DB 379 ETIQGICIPDAPVCTAGKEGCECKCINREPKGCAKPLVWNE-NCKKVCPADK- 436
 QY 237 SQCKTGTGTSTRTVNDNPECHLYKRETRICEVPCGO--PYVSLKKGKCKSKTKSPFVPTYA 301
 DB 437 QMSPPGCGSGSKFNKLTQCCECDGASASGKGLKRMNADTCCECPGMPRECGKQTIWSD 496
 QY 281 KKGKCKSKT-----KKSPEVPTVAGCLSVKKYRP--- 311
 DB 497 KCKCECSPITTCQAPOLLNLTCCCKCPVNMIAKRECKSPROWTDKCLCECSTPATC 556
 QY 312 ---KYCG--SCV---DGRCCPPLTRTYKMRFCEDGET-FSKNVMMIOGSKCNVNC 359
 DB 557 EGRQWCGEAGQCCICPGDKNKGKRPDKPSCCKCKKNPNPTSPQVMDADDEC--KC 614

RESULT 14
 MOCCL_RAT
 ID MOCCL_RAT STANDARD; PRT: 837 AA.
 AC P98089;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Intestinal mucin-like protein (MLP) (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Intestine;
 RX MEDLINE=92184794; PubMed=1371999;
 RA Xu G., Huan L.-D., Khatri I., Wang D., Bennick A., Fahim R.E.F.,
 RT Forstner G.G., Forstner J.F.;
 RL "cDNA for the carboxyl-terminal region of a rat intestinal mucin-like
 peptide.";
 RT J. Biol. Chem. 267:5401-5407(1992).
 CC -1- SUBUNIT: MULTIMERIC.
 CC -1- SUBCELLULAR LOCATION: Secreted.


```
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR003129; TSPN.
DR InterPro: IPR001007; VWF_C.
DR InterPro: IPR003367; tsp_3.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00090; tsp_1; 3.
DR Pfam: PF00093; vwc; 1.
DR Pfam: PF02210; TSPN; 1.
DR Pfam: PF02412; tsp_3; 8.
DR SMART: SM00181; EGF_2.
DR SMART: SM00001; EGF_like; 1.
DR SMART: SM00209; TSP1; 3.
DR SMART: SM00210; TSPN; 1.
DR SMART: SM00214; vwc; 1.
DR PROSITE: PS00022; EGF_1; FALSE_NEG.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00092; TSP1; 3.
DR PROSITE: PS01208; VWF_C; 1.
DR Glycoprotein: Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; signal.
FT SIGNAL 1 22
FT CHAIN 23 1178
FT DOMAIN 7 232
FT DOMAIN 324 381
FT DOMAIN 387 438
FT DOMAIN 441 499
FT DOMAIN 500 553
FT DOMAIN 555 595
FT DOMAIN 596 633
FT DOMAIN 654 698
FT DOMAIN 731 766
FT DOMAIN 767 789
FT DOMAIN 790 825
FT DOMAIN 826 848
FT DOMAIN 849 886
FT DOMAIN 887 922
FT DOMAIN 923 958
FT DOMAIN 959 1178
FT SITE 934 935
FT DISULFID 559 570
FT DISULFID 564 580
FT DISULFID 583 594
FT DISULFID 600 616
FT DISULFID 607 625
FT DISULFID 628 652
FT DISULFID 658 671
FT DISULFID 665 684
FT DISULFID 686 697
FT CARBOHYD 157 157
FT CARBOHYD 244 244
FT CARBOHYD 317 317
FT CARBOHYD 322 322
FT CARBOHYD 463 463
FT CARBOHYD 590 590
FT CARBOHYD 716 716
FT CARBOHYD 1075 1075
SQ SEQUENCE 1178 AA; 131816 MW; F37E02F42C8717A2 CRC64;

Query Match 7.28; Score 152.5; DB 1; Length 1178;
Best Local Similarity 23.04; Pred. No. 0.00052;
Matches 65; Conservative 33; Mismatches 102; Indels 83; Gaps 13;

QY 100 CEYNSRIYNGSEFQPNCKNOCTIDGAVGCIPLCPQELSLPNIGCPNRLVAVTGOCE 159
DB 326 CWDGGRVFADESEWIVDSCTKCTQDSKIYC-----HQITCPVSCADPSFTE--GECCP 378
QY 160 EWWCEDSDSIKPMEDDGLLKEIGFPASEVEELTRNNELIAYGSSILKRLPVFGMEPRI 219
DB 379 --VC---SHSDDESE-----GMSFWSMDTKCSVTGCGTQMRGRSCDVTRSACTG--PHI 426
QY 220 LYNPLGQKC---IVQF-----TSMQCKTKCGTGISTRVYNDNPECRLV----- 261
DB 427 QTRWCSFKKCDHRIRODGCWHSWSPMSSCSVTGCGVGNITRIILCNSPIDPMGKKNVGVNG 486
```

```
QY 262 KENRICEVRPC-----GOPYSSLKKG 283
DB 487 RETEKCEKAPCPVNGMGPMSPMSACTVTCGGGIRERSRLCNSPEEPOYGKPCVGTROH 546
QY 284 KCSKTKKSPEPVRYAGCLSVKRYKPKYCGSCVDGR--CCIP 325
DB 547 DMCKNK--RDCP-----IDGCLSNPCFPFGAFCNCNYPDGSWSCGP 582
```

Search completed: July 8, 2003, 12:29:23
Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2003, 12:22:57 ; Search time 80 Seconds
(without alignments)
981.299 Million cell updates/sec

Title: US-09-901-910-2
Perfect score: 2115
Sequence: 1 MSSRIARALAVTLLHLTR.....ANEAAPFYRLNDHKFRD 381

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1938	91.6	379	11 Q9ES72	Q9ES72 rattus norv
2	1937	91.6	379	11 Q9WTW9	Q9WTW9 rattus norv
3	1667.5	78.8	375	13 Q98TX5	Q98TX5 xenopus lae
4	1596	75.5	334	4 Q9UID7	Q9UID7 homo sapien
5	959	45.3	347	11 Q9RIE9	Q9RIE9 rattus norv
6	957	45.2	347	11 Q9WV51	Q9WV51 rattus norv
7	956.5	45.2	348	11 Q922U0	Q922U0 mus musculu
8	942	44.5	343	13 Q42607	Q42607 xenopus lae
9	934	44.2	344	13 Q98TC8	Q98TC8 gallus gall
10	920.5	43.5	347	13 Q9PT80	Q9PT80 notophthalm
11	882	41.7	349	6 Q97765	Q97765 sus scrofa
12	819.5	38.7	351	11 Q9OZ05	Q9OZ05 rattus norv
13	771	36.5	367	4 Q9S388	Q9S388 homo sapien
14	759	35.9	367	11 Q54775	Q54775 mus musculu
15	759	35.9	367	11 Q99P0	Q99P0 rattus norv
16	628.5	29.7	331	4 Q95958	Q95958 homo sapien

17	615	29.1	354	4 Q95389	Q95389 homo sapien
18	523.5	24.8	250	4 Q76076	Q76076 homo sapien
19	523.5	24.8	251	11 Q920G4	Q920G4 mus musculu
20	495.5	23.4	280	4 Q9HCS3	Q9HCS3 homo sapien
21	491	23.2	250	11 Q9JHC6	Q9JHC6 rattus norv
22	489.5	23.1	176	13 Q9PSS6	Q9PSS6 gallus gall
23	377.5	17.8	128	11 Q9R2C0	Q9R2C0 rattus norv
24	341	16.1	100	4 Q9UDL6	Q9UDL6 homo sapien
25	341	16.1	113	11 Q92164	Q92164 rattus norv
26	336	15.9	119	11 Q920M6	Q920M6 mus splicile
27	336	15.9	119	11 Q91V29	Q91V29 mus musculu
28	293	13.9	230	4 Q8WYK7	Q8WYK7 homo sapien
29	195	9.2	77	4 Q9UDE4	Q9UDE4 homo sapien
30	187	8.8	2327	13 Q9IBG7	Q9IBG7 xenopus lae
31	176	8.3	70	13 Q9DE21	Q9DE21 scyllorhinu
32	174	8.2	179	5 Q9VVR0	Q9VVR0 dtrosophila
33	174	8.2	1045	5 Q8T3A6	Q8T3A6 caenorhabd1
34	174	8.2	1070	5 Q8T3A7	Q8T3A7 caenorhabd1
35	174	8.2	1111	5 Q9XWD6	Q9XWD6 caenorhabd1
36	171.5	8.1	1036	4 Q9NZV1	Q9NZV1 homo sapien
37	171	8.1	1664	5 Q9TVQ2	Q9TVQ2 caenorhabd1
38	170	8.0	1704	5 Q94446	Q94446 chironomus
39	169.5	8.0	1028	11 Q9JDL0	Q9JDL0 mus musculu
40	164	7.8	58	6 Q97574	Q97574 bos taurus
41	162.5	7.7	792	13 Q9OZ43	Q9OZ43 gallus gall
42	162.5	7.7	1095	13 Q9OXG4	Q9OXG4 gallus gall
43	160.5	7.6	2104	5 Q21281	Q21281 caenorhabd1
44	160.5	7.6	2104	5 Q964N4	Q964N4 caenorhabd1
45	159.5	7.5	406	11 Q92513	Q92513 mus musculu

ALIGNMENTS

RESULT 1

ID	Q9ES72	PRELIMINARY	PRT	379 AA.
AC	Q9ES72			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Cysteine-rich protein 61.			
CN	CYR61.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=LUNG;			
RX	MEDLINE=20435857; PubMed=10852911;			
RA	Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,			
RA	Nitsch R.M.;			
RT	*Muscarinic acetylcholine receptors induce the expression of the			
RT	immediate early growth regulatory gene CYR61.*;			
RT	J. Biol. Chem. 275:28929-28936(2000).			
RL	EMBL, AF218568; AAC14964.1.-			
DR	InterPro; IPR000359; Cys_knot.			
DR	InterPro; IPR000867; Tnsl_gro_fac-pr.			
DR	InterPro; IPR000884; TSP1.			
DR	InterPro; IPR001007; VWF_C.			
DR	Pfam; PF00007; Cys_knot.1.			
DR	Pfam; PF000219; IGFBRP.1.			
DR	Pfam; PF00090; TSP_1.1.			
DR	Pfam; PF00094; VWC.1.			
DR	SMART; SM00041; CT.1.			
DR	SMART; SM00121; IB.1.			
DR	SMART; SM00209; TSP1.1.			
DR	SMART; SM00214; VWC.1.			
DR	PROSITE; PS01185; CTCK_1.1.			
DR	PROSITE; PS01225; CTCK_2.1.			
DR	PROSITE; PS00222; IGF_BINDING.1.			
DR	PROSITE; PS01208; VWC.1.			

SQ SEQUENCE 379 AA: 41687 MW: 628F0BBA4C5AFDE9 CRC64;
 Query Match 91.6%; Score 1938; DB 11; Length 379;
 Best Local Similarity 91.4%; Pred. No. 4.2e-178;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSRIARALAVLTLLHRLALSTPCPACHCPLKAPGVGLVRDGGCCCKYCAKOL 60
 1 MSSSTKTILAVATLLHRLALSTPCPACHCPLKAPGVGLVRDGGCCCKYCAKOL 60
 DB 1 MSSSTKTILAVATLLHRLALSTPCPACHCPLKAPGVGLVRDGGCCCKYCAKOL 60
 QY 61 NECCSTQPCDHTKGLKCNFGASTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
 61 NECCSTQPCDHTKGLKCNFGASTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
 DB 61 NECCSTQPCDHTKGLKCNFGASTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDAVAGCIPCLCPDELSPNLGCPNPLVYVGGCCCEWYCDDESIKDPMDODGLLG 180
 121 CTCIDAVAGCIPCLCPDELSPNLGCPNPLVYVGGCCCEWYCDDESIKDPMDODGLLG 180
 DB 121 CTCIDAVAGCIPCLCPDELSPNLGCPNPLVYVGGCCCEWYCDDESIKDPMDODGLLG 180
 QY 181 KEIGFDPASEVELTRNNELIANGKSSLRPLVFGMEPRILYNPL--OGKCIYQTTSMQ 238
 179 --IGFDPASEVELTRNNELIANGKSSLRPLVFGTEPRVLYNPLAHAGKCIYQTTSMQ 236
 DB 179 --IGFDPASEVELTRNNELIANGKSSLRPLVFGTEPRVLYNPLAHAGKCIYQTTSMQ 236
 QY 239 CSKTCGTGISTVYNDNPECRLYKETRICEVRPCGOPYSSLKKGKSKTKKSPBPVR 298
 237 CSKTCGTGISTVYNDNPECRLYKETRICEVRPCGOPYSSLKKGKSKTKKSPBPVR 296
 DB 237 CSKTCGTGISTVYNDNPECRLYKETRICEVRPCGOPYSSLKKGKSKTKKSPBPVR 296
 QY 299 TYAGCLSVKRYPRKYGCGSCVDRCCTPOLRTVYKMFRCDEDETFESKNYMIQSCKNYN 358
 297 TYAGCSVYKRYPRKYGCGSCVDRCCTPOLRTVYKMFRCDEDETFESKNYMIQSCKNYN 356
 DB 297 TYAGCSVYKRYPRKYGCGSCVDRCCTPOLRTVYKMFRCDEDETFESKNYMIQSCKNYN 356
 QY 359 CPHANEAPFPYRLFNDFHFRD 381
 357 CPHANEAPFPYRLFNDFHFRD 379
 DB 357 CPHANEAPFPYRLFNDFHFRD 379

RESULT 2
 Q9WTM9 PRELIMINARY; PRT; 379 AA.
 ID Q9WTM9
 AC Q9WTM9; PRELIMINARY; PRT; 379 AA.
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CYR61 precursor.
 GN CYR61.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=12M; TISSUE=AORTA;
 RA Unoaki H., Yonekura H., Furukawa K., Yamamoto H.;
 RT "Rat Cyr61 mRNA."
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB015877; BAA78339.1; -
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR001007; VWF_C.
 DR InterPro: IPR00007; Cys_knot.1.
 DR Pfam: PF000219; IGBP.1.
 DR Pfam: PF00090; tsp_1.1.
 DR Pfam: PF00093; vwc_1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; VWC.1.
 DR PROSITE: PS01185; CTCK_1.1.
 DR PROSITE: PS01225; CTCK_2.1.
 DR PROSITE: PS00222; IGF_BINDING.1.
 DR PROSITE: PS01208; VWF_C.1.
 KW SIGNAL.
 FT SIGNAL 1 24 POTENTIAL.

FT CHAIN 25 379 CYR61.
 SQ SEQUENCE 379 AA: 41728 MW: D2ABAFD7B84762B CRC64;
 Query Match 91.6%; Score 1937; DB 11; Length 379;
 Best Local Similarity 91.1%; Pred. No. 5.3e-178;
 Matches 349; Conservative 10; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSRIARALAVLTLLHRLALSTPCPACHCPLKAPGVGLVRDGGCCCKYCAKOL 60
 1 MSSSTKTILAVATLLHRLALSTPCPACHCPLKAPGVGLVRDGGCCCKYCAKOL 60
 DB 1 MSSSTKTILAVATLLHRLALSTPCPACHCPLKAPGVGLVRDGGCCCKYCAKOL 60
 QY 61 NECCSTQPCDHTKGLKCNFGASTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
 61 NECCSTQPCDHTKGLKCNFGASTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
 DB 61 NECCSTQPCDHTKGLKCNFGASTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDAVAGCIPCLCPDELSPNLGCPNPLVYVGGCCCEWYCDDESIKDPMDODGLLG 180
 121 CTCIDAVAGCIPCLCPDELSPNLGCPNPLVYVGGCCCEWYCDDESIKDPMDODGLLG 180
 DB 121 CTCIDAVAGCIPCLCPDELSPNLGCPNPLVYVGGCCCEWYCDDESIKDPMDODGLLG 180
 QY 181 KEIGFDPASEVELTRNNELIANGKSSLRPLVFGMEPRILYNPL--OGKCIYQTTSMQ 238
 179 --IGFDPASEVELTRNNELIANGKSSLRPLVFGTEPRVLYNPLAHAGKCIYQTTSMQ 236
 DB 179 --IGFDPASEVELTRNNELIANGKSSLRPLVFGTEPRVLYNPLAHAGKCIYQTTSMQ 236
 QY 239 CSKTCGTGISTVYNDNPECRLYKETRICEVRPCGOPYSSLKKGKSKTKKSPBPVR 298
 237 CSKTCGTGISTVYNDNPECRLYKETRICEVRPCGOPYSSLKKGKSKTKKSPBPVR 296
 DB 237 CSKTCGTGISTVYNDNPECRLYKETRICEVRPCGOPYSSLKKGKSKTKKSPBPVR 296
 QY 299 TYAGCLSVKRYPRKYGCGSCVDRCCTPOLRTVYKMFRCDEDETFESKNYMIQSCKNYN 358
 297 TYAGCSVYKRYPRKYGCGSCVDRCCTPOLRTVYKMFRCDEDETFESKNYMIQSCKNYN 356
 DB 297 TYAGCSVYKRYPRKYGCGSCVDRCCTPOLRTVYKMFRCDEDETFESKNYMIQSCKNYN 356
 QY 359 CPHANEAPFPYRLFNDFHFRD 381
 357 CPHANEAPFPYRLFNDFHFRD 379
 DB 357 CPHANEAPFPYRLFNDFHFRD 379

RESULT 3
 Q98TX5 PRELIMINARY; PRT; 375 AA.
 ID Q98TX5
 AC Q98TX5; PRELIMINARY; PRT; 375 AA.
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Secreted cysteine-rich protein cyr61.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Latinkic B.V., Bennett B., Smith J.C.;
 RA "Characterization of Xenopus cyr61."
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF320592; AAK00947.1; -
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR001007; VWF_C.
 DR InterPro: IPR00007; Cys_knot.1.
 DR Pfam: PF000219; IGBP.1.
 DR Pfam: PF00090; tsp_1.1.
 DR Pfam: PF00093; vwc_1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR SMART: SM00214; VWC.1.
 DR PROSITE: PS01185; CTCK_1.1.
 DR PROSITE: PS01225; CTCK_2.1.
 DR PROSITE: PS00222; IGF_BINDING.1.
 DR PROSITE: PS01208; VWF_C.1.
 KW SIGNAL.
 FT SIGNAL 375 AA: 41460 MW: 78075CA7B380304E CRC64;

Query Match	Similarity	78.8%	Score 1667.5	DB 13	Length 375
Best Local	Similarity	79.6%	Pred. No. 4,3e-152		
Matches	Conservative	28	Mismatches	41	Indels
					Gaps
					5
QY	11	LVVTLH-LTRIALSTPCPACHCPLAEKRCAPGVGLVNDGCGCCRCVAKQLNEDCSKTP	69		
		1 : 11	11 : 11	11 : 11	11 : 11
Db	10	LAIALSGFIDLAIVSSPCAPVCCQPEYVKRCAGVGLVLDGCGCCRICAKQLNEDCSKTP	69		
QY	70	CDHRKGLCECNCGASTALKGTICRAOSBGRPCGYNSRITVONGSFPQNCNHOCTCIDGANG	129		
		1 : 11	11 : 11	11 : 11	11 : 11
Db	70	CDHRKGLCECNCGASSRAIKGTICRAKSEBRPCYNSKITYONGSEFPNCHHOCTCIDGANG	129		
QY	130	CIPLCPEQLSPNLGCPNDRIVKYVNGGCCCEWVGCEDSTIKDMEPODGLLKGELGPASE	189		
		1 : 11	11 : 11	11 : 11	11 : 11
Db	130	CLPCLPQSLSPNLGCPNDRIVKYVNGGCCCEWVGCEDSTIKDMEPODGLLKGELGPASE	187		
QY	190	VELTRNNELIIVGKSSLSKRLPVGFMERILLYNPLOGOKCIVOTTWSQCSKTCGTGIST	249		
		1 : 11	11 : 11	11 : 11	11 : 11
Db	188	GELTRKNEFEVAVIKG-GLKMLPVFSDSP--SHVENSNCIAYOTTWSQCSKTCGTGIST	244		
QY	250	RVTDNDNPCRIVYKEGRICEVVRPCGGPVYSLSLKGKSKSTKKSPEVVRFTYVAGCLSVYK	309		
		1 : 11	11 : 11	11 : 11	11 : 11
Db	245	RVTDNDSNCRIVREFRICEVVRPCGGPSTSLKGKSKCTYTKKSQAARFTYVAGCSSVYK	304		
QY	310	RPKYCGSCVDRCCCTPOLRTYVKMFRCEDEGETFEKKNVMMIOSCACNTYNCNPAHEAAFP	369		
		1 : 11	11 : 11	11 : 11	11 : 11
Db	305	RPKYCGSCVDRCCCTPQOTRTYKIFRCDEDEGETFKKNVMMIOSCACNTYNCNPHNE-AYRY	363		
QY	370	YRLFNDIHKFRD 381			
		1 : 11	11 : 11	11 : 11	11 : 11
Db	364	YRLFNDIHKFRD 375			

RESULT	4
09UID7	
ID	09UID7 PRELIMINARY; PRT; 334 AA.
AC	Q9UID7;
DT	- 01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	CYR61 protein..
GN	CYR61.
OS	'Homo sapiens (Human)'.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCL_TaxID=9606;
RN	[1]
RP	SEQUENCE OF 1-107 FROM N.A.
RC	TISSUE-KIDNEY;
RA	Anding B., Long Y.;
RT	"Cloning of a new gene down-regulated in the small-cell tumor embryonal-rhabdomyosarcoma (RMS).";
RL	Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF003114; AAF21597.1; -
DR	InterPro; IPR000359; Cys_Knot.
DR	InterPro; IPR000867; Insl_gro_fac_pr.
DR	InterPro; IPR000884; TSP1.
DR	InterPro; IPR001007; VWF_C.
DR	Pfam; PF00007; Cys_Knot; 1.
DR	Pfam; PF00219; IGFBP_1.
DR	Pfam; PF00090; tsp_1; 1.
DR	Pfam; PF00093; wvc; 1.
DR	SMART; SM00041; CT; 1.
DR	SMART; SM00209; TSP1_1.
DR	SMART; SM00214; WVC; 1.
DR	PROSITE; PS01225; CTCCK_2; 1.
DR	PROSITE; PS01208; VWFC; 1.
SQ	SEQUENCE 334 AA; 37246 MW; 9188987A7352E948 CRC64;
Query Match	75.5%; Score 1596; DB 4; Length 334;
Best Local Similarity	86.3%; Pred. No. 2,8e+145;
Matches 297; Conservative	4; Mismatches 21; Indels 22; Gaps 4;

QY	49	GGCGCCVCAKQJLNEDSCSKQPCDHNKKGLECNNGASSTLKICICAOSEGRCEVNSIYO	108
Db	2	GGCTHBNLNCILHGHRTASPTSYKHHTKGTECENNGASTLKKICRAOSEGRCEVNSIYO	61
QY	109	NGESFOPNCKHOCTCIDGAVGCIPLCPQPELSIPNLGCPNPLRYVTCOCCSEWVCDSDSI	168
Db	62	NGESFOPNCKHOCTCIDGAVGCIPLCPQPELSIPNLGCPNPLRYVTCOCCSEWVCDSDSI	121
QY	169	KDEMEDODGILGKEIGFDSAEVELTRNNELIAVGSSLSKLPLVFGMEPRILYNPLQOK	228
Db	122	KDPMEDODGILGKEIGFDSAEVELTRNNELIAVGSSLSKLPLVFGMEPRILRYNPLQOK	181
QY	229	CIYQTTSMWSOCSKTGCTGISTRYVNDNPECLRYKETRICERYPCGQOPYSLKGGKCSK	288
Db	182	CIYQTTSMWSOCSKTGCTGISTRYVNDNPECLRYKETRICERYPCGQOPYSLKGGKCSK	241
QY	289	TKKSPEVRYTYAGCLSYKKYRPKYGSGSYDGRCTPOLJTYVYKMRFCED-----GE	341
Db	242	TKKSPEVRYTYAGCLSYKKYRPKYGSGSYDGRCTPOLJTYVYKMRFCED-----GE	295
QY	342	TFESKN-----VMQJQSCKNYVNCPHANEAFPRYLRFNDIHKFRD	381
Db	296	IFQBRHDDPYL-----KCNINCPHANEAFPRYLRFNDIHKFRD	334

RESULT 5			
ID	Q9RIE9	PRELIMINARY:	PRT: 347 AA.
AC	Q9RIE9		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)	
DE	Connective tissue growth factor.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_Taxid=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20145935; PubMed=10679821;		
RA	Xu J., Smock S.L., Saladi F.F., Rosenzweig A.B., Odgren P.R.,		
RA	Marks S.C. Jr., Owen T.A., Popoff S.N.;		
RT	"Cloning the full-length cDNA for rat connective tissue growth factor:		
RT	implications for skeletal development.";		
RL	J. Cell. Biochem. 77:103-115(2000).		
DR	EMBL: AF120275; AAD39132.1; -		
DR	InterPro: IPR000359; Cys_knot.		
DR	InterPro: IPR000867; Insl_9ro_fac_pr.		
DR	InterPro: IPR000884; TSP1.		
DR	InterPro: IPR001007; WVF_C.		
DR	Pfam: PF000007; Cys_knot. 1.		
DR	Pfam: PF00219; IGFBP. 1.		
DR	Pfam: PF00090; tsp_1; 1.		
DR	Pfam: PF00093; WVC. 1.		
DR	SMART: SM00041; CT. 1.		
DR	SMART: SM00121; IB. 1.		
DR	SMART: SM00209; TSP1. 1.		
DR	SMART: SM00214; WVC. 1.		
DR	PROSITE: PS01185; CRCK_1; 1.		
DR	PROSITE: PS01225; CRCK_2; 1.		
DR	PROSITE: PS00222; IGF_BINDING; 1.		
DR	PROSITE: PS01208; WVC; 1.		
SO	SEQUENCE 347 AA; 37756 MW; CFEELAI9766B7B16 CRC64;		

	Query Match	45.3%	Score 959	DB 11	length 347;
	Best Local Similarity	46.5%	Score No. 5	3e-84	
	Matches	177;	Conservative	57;	Mismatches 107; Indels 40; Gaps 7
OY	1	MSRRARALAVTLLHLTRAL-STCPACACPIE-APKARPGVGLVDGGCCCKVAK	58		
		..:::.....	11		
Db	1	MTASVAGYSTALVTLLCTCRPATGDCSCQCCCAAAAFRCNAGSVLVDGGGCCCKVAK	60		
OY	59	OLNEDCSKTPQPCDHTKCECNEGASSTALKGICRAOSBGRPCENSRIVQNCSFOPNCK	118		

```
Db      61 QUGELCTERDPCDHPKGLFCDFGSPANKRIGVCTAK-DGAPCVFGGSYRSGESFOSOCK 119
Qy      119 HOCSTIDGAVGCIPLCPQELSLPNLGCPNRLVKTGTGOCCEWVDEDSIKDPMEDDGL 178
Db      120 YOCCTLDGAVGCVPLCSMDVRLPSDCCPPRRVKLPKGCCEWVDEP----- 167
Qy      179 LKELGFPASSEVELTRNNELLAVKSSILKRL-PVFGMEPRILYNPLGOKCIYOTTSMS 237
Db      168 -----KDRYVGPALAAVRLIEDTFEGDPPTM-----RANCLVOTTEWS 205
Qy      238 QCSKTGCTGISTRVYNDNDEPCRLVKEIRICEVRPCGOPYSSLLKGGKCSKTGSPPEVR 297
Db      206 ACSKTGCGMISTRVYNDNDEPCRLVKEIRICEVRPCGOPYSSLLKGGKCSKTGSPPEVR 265
Qy      298 FTYACCLSVKKRYPKPYGSCVDRCCTPOLRTVYMRRCEDGETFSNMMIQCCKNY 357
Db      266 FELSCCTSVKTYRAKFCVCTDRCCTPHRTTLTVEFRCPDGEIMKKNMMEIKTCACHY 325
Qy      358 NCPHANEAFPPY--RLFNDI 376
Db      326 NCPGNDIFESLYRKMYGDM 346
```

RESULT 6

Q9WVS1 PRELIMINARY: PRT: 347 AA.

```
AC 09WVS1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Connective tissue growth factor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Tezuka K., Tamatani T.;
RT "Rattus norvegicus connective tissue growth factor.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023068; BAA82125.1; -
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00007; VWF_C.
DR Pfam; PF00007; Cys_knot.1.
DR Pfam; PF00019; IGFBP.1.
DR Pfam; PF00090; tsp_1.1.
DR Pfam; PF00093; vwc.1.
DR SMART; SM00041; CT.1.
DR SMART; SM00121; IB.1.
DR SMART; SM00209; TSP1.1.
DR SMART; SM00214; VWC.1.
DR PROSITE; PS01185; CTCK_1.1.
DR PROSITE; PS01225; CTCK_2.1.
DR PROSITE; PS00222; IGF_BINDING.1.
DR PROSITE; PS01208; VWF_C.1.
SQ SEQUENCE 347 AA; 37837 MW; 6A69511DE72FEF1C CRC64;
```

Query Match 45.2%; Score 957; DB 11; Length 347;
Best Local Similarity 46.5%; Pred. No. 8.2e-84;
Matches 177; Conservative 57; Mismatches 107; Indels 40; Gaps 7;

```
Qy      1 MSSRIARALAVITLHLTRLAL-STCPACHCPLF-APKAPGVGLVRDCCGCCRYAK 58
Db      1 MLASVAGPISLALVILCTRPATGDCSAQCQCAEAAPCPAGVSLVLDGCGCCRYAK 60
Qy      59 QLNEDCSKTOPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNC 118
Db      61 QUGELCTERDPCDHPKGLFCDFGSPANKRIGVCTAK-DGAPCVFGGSYRSGESFOSOCK 119
Qy      119 HOCSTIDGAVGCIPLCPQELSLPNLGCPNRLVKTGTGOCCEWVDEDSIKDPMEDDGL 178
```

```
Db      120 YOCCTLDGAVGCVPLCSMDVRLPSDCCPPRRVKLPKGCCEWVDEP----- 167
Qy      179 LKELGFPASSEVELTRNNELLAVKSSILKRL-PVFGMEPRILYNPLGOKCIYOTTSMS 237
Db      168 -----KDRYVGPALAAVRLIEDTFEGDPPTM-----RANCLVOTTEWS 205
Qy      238 QCSKTGCTGISTRVYNDNDEPCRLVKEIRICEVRPCGOPYSSLLKGGKCSKTGSPPEVR 297
Db      206 ACSKTGCGMISTRVYNDNDEPCRLVKEIRICEVRPCGOPYSSLLKGGKCSKTGSPPEVR 265
Qy      298 FTYACCLSVKKRYPKPYGSCVDRCCTPOLRTVYMRRCEDGETFSNMMIQCCKNY 357
Db      266 FELSCCTSVKTYRAKFCVCTDRCCTPHRTTLTVEFRCPDGEIMKKNMMEIKTCACHY 325
Qy      358 NCPHANEAFPPY--RLFNDI 376
Db      326 NCPGNDIFESLYRKMYGDM 346
```

RESULT 7

Q922U0 PRELIMINARY: PRT: 348 AA.

```
AC 0922U0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Connective tissue growth factor.
GN CTGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strusberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006783; AAH06783.1; -.
DR MGI; MGI:95537; Ctgf.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00007; VWF_C.
DR Pfam; PF00019; IGFBP.1.
DR Pfam; PF00090; tsp_1.1.
DR PROSITE; PS01185; CTCK_1.1.
DR PROSITE; PS01225; CTCK_2.1.
DR PROSITE; PS00222; IGF_BINDING; UNKNOWN_1.
DR PROSITE; PS01208; VWF_C; UNKNOWN_1.
SQ SEQUENCE 348 AA; 37794 MW; 4D7BD9089174049 CRC64;
```

Query Match 45.2%; Score 956.5; DB 11; Length 348;
Best Local Similarity 46.9%; Pred. No. 9.2e-84;
Matches 179; Conservative 56; Mismatches 106; Indels 41; Gaps 8;

```
Qy      1 MSSRIARALAVITLHLTRLAL-STCPACHCPLF-APKAPGVGLVRDCCGCCRYAK 57
Db      1 MLASVAGPISLALVILCTRPATGDCSAQCQCAEAAPCPAGVSLVLDGCGCCRYAK 60
Qy      58 QLNEDCSKTOPCDHTKGLCECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNC 117
Db      61 QUGELCTERDPCDHPKGLFCDFGSPANKRIGVCTAK-DGAPCVFGGSYRSGESFOSOCK 119
Qy      118 HOCSTIDGAVGCIPLCPQELSLPNLGCPNRLVKTGTGOCCEWVDEDSIKDPMEDDGL 177
Db      120 YOCCTLDGAVGCVPLCSMDVRLPSDCCPPRRVKLPKGCCEWVDEP----- 168
Qy      178 LKELGFPASSEVELTRNNELLAVKSSILKRL-PVFGMEPRILYNPLGOKCIYOTTSMS 236
Db      169 -----KDRYVGPALAAVRLIEDTFEGDPPTM-----RANCLVOTTEWS 205
```


OY	23	SQSKTGCGISIRATNDNDRPECLYKERRICEVPRCQGVYSLSLAKKKCKSTTKSPFV	296
	206	SASCSCTGAGISTRTANDTFCRLRKOSRSLCMVPCEADLEINLKKKKCICRTPKLAPV	265
OY	297	RFTYACGLSVKKRYRPKYGSCVDGRCPQLTRTYVKMRECEDEFTSGKNVMQSQCN	356
		: :	
Dd	266	KFLSLSCTSVKTIYRAAFGCVGCTDGRKCTPHRTITLVFEKCPDGEIMKKMMIFKTCAH	325
OY	357	YNCPHANEAAPFY--RLFNDI	376
Dd	326	YNCPGDNDIPESLYIKRMKGDM	347

RESULT 8

```

ID      042607          PRELIMINARY;          PRT;       343 AA.
AC      042607:
DT      01-JAN-1998 (TrEMBLrel. 05, Created)
DT      01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Connective tissue growth factor XCTGF.
DE      Xenopus laevis (African clawed frog).
OC      Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Xenopodinae; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxID=8355;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Ying Z., King M.L.;
RL      Submitted (AUG-1997) to the EMBL/genbank/dbj databases.
RS      EMBL: U43524; AAB67638.1; --
DR      EMBL: U43523; AAB67638.1; --
DR      InterPro: IPR000359; Cys_knot.
DR      InterPro: IPR000867; Insl_gro_fac.
DR      InterPro: IPR000884; TSP1.
DR      InterPro: IPR001007; WVF_C.
DR      Pfam: PF00007; Cys_knot. 1.
DR      Pfam: PF00219; IGFBP. 1.
DR      Pfam: PF00090; tsp_1. 1.
DR      Pfam: PF00093; wvc. 1.
DR      SMART: SM00041; CT. 1.
DR      SMART: SM00121; IB. 1.
DR      SMART: SM00209; TSP1. 1.
DR      SMART: SM00214; WVC. 1.
DR      PROSITE: PS01185; CTCK_1. 1.
DR      PROSITE: PS01225; CTCK_2. 1.
DR      PROSITE: PS01208; WVEC. 1.
SO      SEQUENCE   343 AA;  37966 MW;   93E221C5DB565A81 CRC64;

```

Query Match	44.58;	Score 942;	DB 13;	Length 343;
Best Local Similarity	47.58;	Pred. No. 2.2e-82;		
Matches 169;	Conservative 50;	Mismatches 95;	Indels 42;	Gaps 6

[illegible]

QY 323 CTPLTTRTVKMFRCSEDETFESKVVMMIOSCKNVCNCPHANE--AATPEYRLENDI 376
||| | : : | : ||| || : | : ||| | : : : : | :
Db 287 CTPIHRAITPLVEEKCSPDGEVMMKKMMMFITCCACHENCSPGNDIIFEAANYRKKMGDM 3422

RESULT 9

ID	098708	PRELIMINARY;	PRT;	344 AA.
AC	098708:			
DT	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Connective tissue growth factor precursor (Connective tissue growth factor/hypertrophic chondrocyte-specific protein 24).			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Gyl D.;			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			

RA	Mutudatay, Kubota S., Takigawa M.;
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL, AJ298335; CAC33438.1; -
DR	EMBL, AF463517; AAL68834.1; -
DR	InterPro; IPR000359; Cys_Knot.
DR	InterPro; IPR000867; Ins1_gro_fac_pr.
DR	InterPro; IPR000884; TSP1.
DR	InterPro; IPR001007; WVF_C.
DR	Pfam; PF00007; Cys_knot; 1.
DR	Pfam; PF00219; IGFBP; 1.
DR	Pfam; PF00090; tsp_1; 1.
DR	Pfam; PF00093; wvc; 1.
DR	SMART; SM00041; CT; 1.
DR	SMART; SM00121; IB; 1.
DR	SMART; SM00209; TSP1; 1.
DR	PROSITE; PS01185; CYCK_1; 1.
DR	PROSITE; PS01225; CYCK_2; 1.
DR	PROSITE; PS00222; IGF_BINDING; 1.
DR	PROSITE; PS01208; WVF; 1.
KW	Signal.
FT	Signal.
FT	CHAIN
SEQUENCE	22 1 21
SEQUENCE	344 AA; 37499 MM; POTENTIAL.
SEQUENCE	344 AA; 37499 MM; CONNECTIVE TISSUE GROWTH FACTOR.
SEQUENCE	344 AA; 37499 MM; 69E639AF69BFD00 CRC64;

Query Match	44.28;	Score 934;	DB 13;	Length 344;
Best Local Similarity	45.68;	Pred. No. 1.3e-81;		
Matches 170; Conservative	58;	Mismatches 105;	Indels 40;	Gaps 7;

[illegible]

```

OY 306 VAKYPRKXGSCVDRCRCCPOLTRFYVKMFPRBEDGTFKFNWMIQSCKNYCNCPHANEA 365
DB 271 VKTYAKRCGCVGTDRCCCPHPTATLPVEFKCPDGEIMRKMMFITYCACHTYCPDNDI 330
OY 366 AFPEY--RLFNDI 376
DB 331 FESLYRRMYGDM 343

RESULT 10
O9PT80 PRELIMINARY; PRT; 347 AA.
AC O9PT80;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Connective tissue growth factor.
CN CTGF.
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC Notophthalmus.
OK NCBI_TaxId=8316;
RN NM
RP SEQUENCE FROM N.A.
RC TISSUE=FORELIMB BLASTEMA;
RA "Gates P.B.: to the EMBL/genbank/DDbJ databases.
RN Submitted (JAN-2000) to the EMBL/genbank/DDbJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=FORELIMB BLASTEMA;
RA "MEDLINE=99033008: PubMed=9813273;
RT "Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;
RT "Identification of newt connective tissue growth factor as a target of
RT retinoid regulation in limb blastenal cells.";
RL Gene 222:119-124(1998).
DR EMBL: AJ271167; CAB65965.1; -.
DR InterPro: IPR000359; Cys_knot.
DR InterPro: IPR000867; Insl_gro_fac-pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WMF_C.
DR Pfam: PF00007; Cys_knot. 1.
DR Pfam: PF00219; IGFBP. 1.
DR Pfam: PF00090; tsp_1. 1.
DR Pfam: PF00093; vwc_1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1. 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS01208; WMFC; UNKNOWN; 1.
SQ SEQUENCE 347 AA; 38098 MW; 3BE2399F27672C1 CRC64;

Query Match 43.5%; Score 920.5; DB 13; Length 347;
Best Local Similarity 45.4%; Pred.No.2.7e-80;
Matches 171; Conservative 48; Mismatches 107; Indels 51; Gaps 6

OY 9 LALVVTLLHTRLTALSTCPAACHCPLEAPKCAPGVGLVDDGGCCCVCAKOLNEDCSKTQ 68
DB 12 LILAVALLSWSCA-QDCSGEGRCKNRKPRECAAGISLVWDGGCCCKVCAKOLGELCTEKD 70
OY 69 PCDHTHGLECNFGASSTALKGICRAOSSEGRPCENYSRTIQNGSEFPQNCCKHOCTCIDGAV 128
DB 71 VCDPIHRLGFLCDRGSVNNKIKIGVCTPAK-DSAPCVFGVMYVRSGESPQSSCKYQCTLDGGV 129
OY 129 GCIPLCPEBELSLPNIGCENPLRYKVTGQCCCEWVDEDSIKPMEDDGGLKELGFDAS 188
DB 130 GCVPICAGVDVRLPSDCFPFRKVKOLGRCCEWVDCDFP----- 168
OY 189 EVELTRNNELIAY-----GKGSLKRLPVPVGMPEPRILIYNPLQGOKTIVOTTSMSQSK 241
DB 189 EVELTRNNELIAY-----GKGSLKRLPVPVGMPEPRILIYNPLQGOKTIVOTTSMSQSK 241

```

Db	169	-EQRRVGPALAVYHQEETVYGPDDSLMR-----	ANCLVQTTEWSACSK	209
Qy	242	TCGCTGSTRVRYNDANEPCRLVYKTRTCEIRAPCCQPIVYSSLKAKKCKSKTKKSPKPAFYTA	301	
Db	210	TCGCMGISTRVRYNDNEMCRLEKQSRCLCMVPCADLEENIKKQKCIIRPKISKIPVKEFLS	269	
Qy	302	GCLSAKKRKRPYKSCGCVDRGCRCTPOLTRIVYKRRFCEDEFTESKNVMYMQSCKNVNCNH	361	
Db	270	GCTSKYKTRAFKFCGVCYTDGRCTTPHRTATLPVEFKCPDGEVKKKKMMF1KTCACHNCFG	329	
Qy	362	ANEAAPFPY--RLFNDI 376		
Db	330	DNDIFESMYRRKMGDM 346		
RESULT 11				
ID	097765	PRELIMINARY;	PRT;	349 AA.
AC	097765;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Connective tissue growth factor.			
GN	CTGF.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.			
OX	NCBI_taxid=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=UTERUS.			
RA	Harding P.A., Briggsstock D.R.;			
RT	"Cloning and sequencing of a porcine connective tissue growth factor			
RT	(CTGF) cDNA."			
RL	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: U70060; AAD00174.1; "			
DR	InterPro: IPR000359; Cys_knot.			
DR	InterPro: IPR000867; Insl_gro_fac-pr.			
DR	InterPro: IPR000884; TSP1.			
DR	InterPro: IPR001007; WME_C.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00219; IGFBP; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; wvc; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00121; IB; 1.			
DR	SMART: SM00209; TSP1; 1.			
DR	SMART: SM00214; WVC; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			
DR	PROSITE: PS01225; CTCK_2; 1.			
DR	PROSITE: PS00222; IGF_BINDING; 1.			
DR	PROSITE: PS01208; WVEC; 1.			
DR	SEQUENCE 349 AA; 35AB4275AC1D4B3A CRC64;			
Query Match 41.7%; Score 882; DB 6; Length 349;				
Best local Similarity 43.5%; Pred. No. 1,3e-76;				
Matches 162; Conservative 59; Mismatches 111; Indels 40; Gaps 7				
Qy	10	ALVYTLTLTRIAL-STCPACHCPL-DAPKAPGVGLVRDGCQCGCKVCAKOLNEDCSKT	67	
Db	12	AFVLLALICSRPASQSDGSCGQCAAGARRACPAVSSILBEGCGCRILCAKILGDLCTTR	71	
Qy	68	QPCDHKTGLKCNFGASSTALGICRAQSGEGRCEYNRSRIYONGSEFQPNCKHQTCTIDGA	127	
Db	72	APCPRHKLGLFDFGSPANRKIGVCTAK-DCAPCVGVTYRSGESFQSCKYQCTCLDGA	130	
Qy	128	VGCIPLPCQELSLPRIGCPNRLVAVYTGOCCEBWCDESDTDPMDODGLLGLKELGPA	187	
Db	131	VGCVPLCSMDVRLPSPPDPRFRRYVLPGKCCBWCDEP-----	169	
Qy	188	SEVELTRNNELIAYGKSSLRKL-PVFGMEPRILYNLPQGGKCIQVOTTSMSQCSKTCGTG	246	
Db	170	-----KDTVVGPPALAAVRLDTFGPDPTMM-----RANCLVQTEWSACSKTCCGNG	216	

```

QY 247 ISRYTNDNPECRVYKTRICRCEVRPGQPVYSSLSKKGKCKSKTKSPPEVRYTACLSY 306
DB 217 ISRYTNDNASCLEKQSRKLCVRCPEADLEENIKKGRKIRPKISKPKELSGCTSY 276
QY 307 KKYRPGYCGSCVDGRCCTPOLTRTYVMRPFCEDEFFSKNVMATIOSCKNYPNANEAA 366
DB 277 KTYRAAFGCVGTJDRCTPHRTTTLPEYFRKCPDGEVYMKKSMMFKTKACHYNGPNDLF 336
QY 367 FPFY--RLFNDI 376
DB 337 ESLYYRKMAYGDM 348

RESULT 12
ID 090205 PRELIMINARY; PRT; 351 AA.
AC 090205;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE NOV protein.
GN NOV.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=20035752; PubMed=10570975;
RA Liu C., Liu X.J., Crowe P.D., Kelnar G.S., Fan J., Barry G., Manu F.,
RA "Nephroblastoma overexpressed gene (NOV) codes for a growth factor
RT that induces protein tyrosine phosphorylation.";
RL EMBL: AF171936; AAD49371.1; -.
DR InterPro: IPR000359; Cys_knot.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp.1; 1.
DR Pfam: PF00093; vwc.1; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS01208; VWF; UNKNOWN_1.
SQ SEQUENCE 351 AA; 38509 MW; 02619707DE7C1BFB CRC64;

Query Match 38.7%; Score 819.5; DB 11; Length 351;
Best Local Similarity 43.4%; Pred. No. 1.4e-70;
Matches 158; Conservative 50; Mismatches 117; Indels 39; Gaps 8;

QY 7 RALATVYTLHLRLALST--CPAC--HCPLEAPKAPGVGLVRDSCGCKCAKOLNE 62
DB 8 QCCCLGFLHLHLINQVATLRPCSPQSPISPCAPGVRSYLDGSCCCPYCARNGE 67
QY 63 DCKSTPCHTKLECNFASSTALGICRAOSEGRPEYNSRIYONGESFQPNCKHCT 122
DB 68 SCSEMPPCQSSGLYCDRSADPNNEGICMY-PEGDNCYVDGVIYRNGEKEFEPNCOYHCT 126
QY 123 CIGANGCPLPCPOELSLPMLGCPNRLVYVGTGCEWYCDSDSIRDPEDDGLGKE 182
DB 127 CRGGQIGCVPRQULDLPLPGPCAPAKVAVPGSCCKWTCG-----SEEGTIG-G 177
QY 183 LGFDASEVELLRNNELIANGKSSLRKPLVFGMEPRILVNPLOGKCIYQTSWQCKT 242

```

```

DB 178 LALPAYRENTAVGVEL-----SDSSI-----NCIEQTEMASCKS 213
QY 243 CGGTGSTRYTNNDPECRVYKTRICRCEVRPGQ--PVYSSLSKKGKCKSKTKSPPEVRYT 301
DB 214 CGMGLSTRYTNRLQCEMKQTRLCVRCPEDEPGATDMKCKKCRKSKLSHLOFK 273
QY 302 GCISVKKYRPGYCGSCVDGRCCTPOLTRTYVMRPFCEDEFFSKNVMATIOSCKNYPN 361
DB 274 NCTSLYTYRPFRCGICSDGRCTPHRTTTLPEYFRKCPDGEVYMKKSMMFKTKACH 333
QY 362 ANEA 365
DB 334 NNEA 337

RESULT 13
ID 095388 PRELIMINARY; PRT; 367 AA.
AC 095388;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Connective tissue growth factor related protein WISP-1.
GN WISP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADULT LUNG, AND FETAL KIDNEY;
RX MEDLINE=99061933; PubMed=9843955;
RA Pennica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Bush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WISP genes are members of the connective tissue growth factor family
RT that are up-regulated in wnt-1-transformed cells and aberrantly
RT expressed in human colon tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Blechschmidt K., Kalaydjieva L., Goodman R., Gresham D., Baas F.,
RA Jonge R.d., Schlibabel M., Schattevoy R., Dette M., Menzel U.,
RA Rosenthal A.;
RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF100779; AAC96321.1; -.
DR EMBL: AF192304; AAF22341.1; -.
DR InterPro: IPR000359; Cys_knot.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp.1; 1.
DR Pfam: PF00093; vwc.1; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS01208; VWF; UNKNOWN_1.
SQ SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;

Query Match 36.5%; Score 771; DB 4; Length 367;
Best Local Similarity 38.3%; Pred. No. 6.7e-66;
Matches 149; Conservative 61; Mismatches 125; Indels 54; Gaps 9;

QY 2 SSRIARLALVYTLHLRLALST-----CPAC--HCPLEAPKAPGVGLVRDSCGCK 54
DB 18 STVALATLSPAPPTMDPTPAILEDTSRQPFCKWPCPCPPSPRCPLGVSLITDGCECK 77

```

```

QY 55 VCAKQUNEDCKSTQPCDHTKGLGECNF--GASSTALGICRAOSEGRPCYNSRIYONGESE 113
DB 78 MCAOQIGDNCTEAAICDHRGLYCDYSGDRPRYAIGVC--AQYVGCVLDGVRYNNGQSF 136
QY 114 QPNCHOCCTIDGAVGCTPILCPQELSLPILGCPNRLVKYVTCGCCWYCDSDSIKDPME 173
DB 137 QPNCHYNTCTCIDGAVGCTPLC-LRVPRPLMCPHRRRISIPHCCEQWYCEDDA-KRPKK 194
QY 174 DODGLGKELGFD--SEVELTRNNELIYAVGKSSILKRLPVFGEPRILYNPLQOKCTVQ 232
DB 195 TAPRDTG---AFDADGVEVAMHRN-----CIAY 219
QY 223 TTSNOCCTGCTGTSTRTVNDNPECRLYKETRICEVRPCGOPYSSLKKGKCKSKTKS 292
DB 220 TSPMPCSTSCGLGYSTRISNVNAQCWPEBSRLCNLRPCVDVDTHTLTKAGKCKLAVYQP 279
QY 293 PEVFTYAGCLSVYKRYKYGSCVDGCTPDLTRVYKMFREDEGTFFKNVMIO 352
DB 280 EASMFETLAGCISTSTYQPKYGVCMNRCCTIPYKSTIDVSFCQPDGLGFSRQYLMWNA 339
QY 353 CKCNVNCPHANEAFPRYLPNDIKFRD 381
DB 340 CFCNLSCRNPND-----IFADLESYPD 361

RESULT 14
054775 PRELIMINARY: PRT; 367 AA.
AC 054775;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ELM1.
GN ELM1 OR WISPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-HEN;
RX MEDLINE=98119879; PubMed=9449709;
RA Shiroishi T., Toma H., Yokota J.;
RA Shiroishi T., Toma H., Yokota J.;
RT "Expression of the Elmi gene, a novel gene of the CCN (connective
RT tissue growth factor, Cyr61/Ce10, and neuroblastoma overexpressed
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735
RT murine melanoma cells."
RT J. Exp. Med. 187:289-296(1998).
RN 12
RP SEQUENCE FROM N.A.
RC TISSUE-MAMMARY;
RX MEDLINE=99061933; PubMed=9843955;
RA Penica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Deuel B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WSP genes are members of the connective tissue growth factor family
RT that are up-regulated in wt-1-transformed cells and aberrantly
RT expressed in human colon tumors."
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
DR EMBL; AB004873; BAA24949.1; -
DR MGD; MGI:1197008; WISPL.
DR InterPro; IPR000359; Cys_Knot.
DR InterPro; IPR000867; Insl_gro_fac-pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; WVF-C.
DR Pfam; PF00007; Cys_Knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; wvc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00041; CT; 1.

```

```

DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; WVEC; UNKNOWN_1.
SQ SEQUENCE 367 AA; 40702 MW; 387C0569EFA5E96 CRC64;

Query Match 35.9%; Score 759; DB 11; Length 367;
Best Local Similarity 38.5%; Pred. No. 9, 6e-65;
Matches 149; Conservative 52; Mismatches 130; Indels 56; Gaps 8;

QY 5 IARALAVTLIHLRLALST-----CPAACHPLEAPKAPGVGLVDGCGCKYCA 57
DB 21 IATLSPTPTTFTFPADLEETTRPPECKWPCDECPSPRCLGVSTLTGCECKICA 80
QY 58 KOLNEDCKSTQPCDHTKGLGECNF--GASSTALGICRAOSEGRPCYNSRIYONGESFQPN 116
DB 81 QQLGDNCTEAAICDHRGLYCDYSGDRPRYAIGVC--AQYVGCVLDGVRYNNGQSFQPN 139
QY 117 CKHOCCTIDGAVGCTPILC--PQELSLPILGCPNRLVKYVTCGCCWYCDSDSIKDPME 174
DB 140 CRYNCTCIDGTYGCTPLDSPRP---PRLMCHQPRHVRVPGCCGQWYCDSDAR---PR 193
QY 175 QDGLGKELGFDSEVELTRNNELIYAVGKSSILKRLPVFGEPRILYNPLQOKCTVDT 234
DB 194 QTALLDTRAFPAASGAVEQRYEN-----CIAYTS 221
QY 235 SMSQCKTQCTGTSTRTVNDNPECRLYKETRICEVRPCGOPYSSLKKGKCKSKTKSPE 294
DB 222 PMSPOSTGCTGISTRTISNVNAQCWPEBSRLCNLRPCVDVDTHTLTKAGKCKLAVYQPE 281
QY 295 PVFTYAGCLSVYKRYKYGSCVDGCTPDLTRVYKMFREDEGTFFKNVMIO 354
DB 282 ATNFTLACVSTRYTRPRYKYGCTGNRCCTIPYKSTISVDGCPGPGFSRQYLMWNA 341
QY 355 CNVNCPHANEAFPRYLPNDIKFRD 381
DB 342 CFCNLSCRNPND-----IFADLESYPD 361

RESULT 15
0999P0 PRELIMINARY: PRT; 367 AA.
AC 0999P0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20487548; PubMed=11031104;
RA Sleeman M.A., Murlson J.G., Strachan L., Kumble K.D., Glenn M.P.,
RA McGrath A., Bickertstaff P., Grierson A., Havukkala I., Tan P.,
RA Watson J.D.;
RT "Gene expression in rat dermal papilla cells: analysis of 2529 ESTs."
RT Genomics 69:214-224(2000).
RL EMBL; AF28049; AAK00729.1; -
DR InterPro; IPR000359; Cys_Knot.
DR InterPro; IPR000867; Insl_gro_fac-pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; WVF-C.
DR Pfam; PF00007; Cys_Knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; wvc; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.

```


THIS PAGE BLANK (USPTO)